



Recherche d'informations dans les systèmes hypertextes : des représentations de la tâche à un modèle de l'activité cognitive

Jean-François Rouet, André Tricot

► To cite this version:

Jean-François Rouet, André Tricot. Recherche d'informations dans les systèmes hypertextes : des représentations de la tâche à un modèle de l'activité cognitive. Sciences et Techniques Educatives, 1995, 2 (3), pp.307-331. edutice-00135871

HAL Id: edutice-00135871

<https://edutice.hal.science/edutice-00135871>

Submitted on 9 Mar 2007

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.

Recherche d'informations dans les systèmes hypertextes : des représentations de la tâche à un modèle de l'activité cognitive

(Information search in hypertext systems: from task representations to a cognitive activity model)

Jean-François Rouet — *Laboratoire Langage et Communication (LACO),
URA 1607 du CNRS, Université de Poitiers, 95 av. du Recteur Pineau, 86 022
Poitiers cedex (e mail : rouet@isis.imag.fr).*

André Tricot — *Centre de Recherche en Psychologie Cognitive (CREPCO),
URA 182 du CNRS, Université de Provence, 29 av. Robert Schuman, 13 621 Aix
en Provence (e mail : raison@romarin.univ-aix.fr).*

Cet article est en partie fondé sur un texte antérieur à paraître dans
H. van Oostendorp (Ed.) Cognitive aspects of electronic text processing. Ablex
Publishing Co.

Titre courant: Recherche d'informations dans les hypertextes

RÉSUMÉ. Dans cet article nous examinons les aspects psychologiques et ergonomiques de la recherche d'informations dans les systèmes hypertextes. Notre hypothèse générale est que la conception de systèmes hypertextes pour l'éducation ne peut faire l'économie d'un modèle cognitif général des tâches et des activités faisant appel à des structures d'informations complexes. Dans une première partie nous évoquons brièvement diverses approches de la notion de tâche de recherche d'informations, et nous proposons une catégorisation en trois niveaux: le modèle rationnel, sa représentation cognitive, et l'activité qui en résulte. Dans la seconde partie nous examinons quelques études empiriques montrant l'influence de facteurs individuels et situationnels sur l'utilisation des hypertextes. Dans la troisième partie, nous proposons le cycle évaluation-sélection-traitement comme ébauche d'un modèle d'activité de recherche d'informations. Nous discutons brièvement les implications possibles de notre approche pour la conception des systèmes hypertextes.

ABSTRACT. In this paper we examine the cognitive and ergonomic aspects of information search in hypertext systems. Our working hypothesis is that the design of hypertext systems for education must rely on a general cognitive model of the tasks and activities concerned

with complex information structures. In the first part we describe briefly several approaches to the notion of information search and we propose a categorization in three levels: Rational task model, cognitive representation, and resulting activity. In the second part we review several studies showing the influence of individual and situational factors on information search in hypertext. In the third part we propose the selection-evaluation-processing cycle as a tentative activity model of information search. We briefly discuss the possible implications of our approach for the design of hypertext systems.

MOTS-CLÉS: hypertexte, lecture, compréhension, recherche d'informations, analyse de tâche, activité cognitive, représentation cognitive, conception, ergonomie, psychologie, interface, organisateurs textuels, apprentissage, utilisabilité.

KEY WORDS: Hypertext, reading, comprehension, information retrieval, task analysis, cognitive activity, cognitive representation, design, human factors, psychology, interface, textual organizers, learning, usability.

1. Introduction

La recherche sur les processus de traitement de l'information s'est orientée ces dernières années vers une prise en compte accrue de situations " finalisées ", dans lesquelles le sujet doit non seulement " traiter " (percevoir, identifier, comprendre) l'information mais encore mettre en oeuvre cette dernière pour accomplir une tâche donnée [AND 90] [HOC 87]. Un secteur de recherche en plein essor actuellement est celui qui concerne l'approche cognitive de l'utilisation de systèmes documentaires complexes, tels les hypertextes. Ces systèmes ont fait l'objet de nombreuses présentations tant en anglais [CON 87] [NIE 90] qu'en français [BAL 90] (voir également [PAS 92]) ce qui rend superflue une description détaillée ici. En revanche, la conception de ces systèmes se heurte toujours à de nombreux obstacles liés à leur mise en oeuvre par les utilisateurs. Un bon indice en est le contraste entre la multiplication des " prototypes " de toutes sortes, et la rareté des applications réellement diffusées et utilisées avec succès. On s'accorde désormais à penser que l'ergonomie des systèmes d'informations est devenue une problématique centrale, au carrefour des sciences cognitives et des sciences de l'information [TRI 95] [WRI 91].

De fait, de nombreuses recherches se sont déjà intéressées aux moyens d'améliorer l'utilisabilité des systèmes hypertextes. Plusieurs approches sont possibles, telles que la modélisation du domaine [NAN 91] ou des tâches [MAR 91], ou encore la mise au point d'outils de navigation de toutes sortes [NIE 90]. Mais ces approches sont parfois limitées, difficiles à appliquer, et ne garantissent pas nécessairement une bonne ergonomie du système. De plus, si la nécessité d'analyser l'utilisateur et ses besoins est reconnue depuis longtemps [THO 89],

encore rares sont les travaux qui prennent effectivement en compte cette dimension [ROU 92].

Dans cet article nous nous proposons de montrer qu'une analyse détaillée des activités de recherche d'informations pourrait guider la conception des systèmes hypertextes. Dans notre perspective, le terme d'hypertexte ne réfère à aucun logiciel particulier, mais plutôt au principe général d'un système d'information non-linéaire, basé sur des "fiches" de texte interconnectées par des liens sémantiques [CON 87]. Quant à l'analyse des tâches, elle consiste d'abord à décrire les types de connaissances procédurales et déclaratives impliqués dans l'activité du sujet. Il faut ensuite modéliser les buts, sous-buts, procédures, stratégies, objets et actions impliqués dans cette activité. Ce type d'approche est largement répandu en ergonomie de la conception [DIA 89] [HAM 93] [JOH 91]. Plusieurs auteurs ont suggéré que l'analyse des tâches pouvait aussi avoir des conséquences positives sur la conception des systèmes d'information. Par exemple, selon [KNI 88],

“ pour déterminer [l'organisation optimale des liens dans un document hypertexte], l'auteur doit prendre en compte les utilisations futures du document (...) Si l'auteur peut anticiper la nature des tâches et fournir des 'cadres' pour assister ces tâches, alors l'utilisateur aura plus de chances de tirer profit du document. ” (p. 339)

Cette approche tout à fait louable se heurte cependant à un sérieux problème : comment "anticiper la nature des tâches" faisant appel à des documents? On ne dispose à l'heure actuelle que de modèles cognitifs généraux de la compréhension en lecture (voir [DEN 90] [DIJ 83]), qui ne concernent qu'indirectement les activités de recherche ou d'utilisation de l'information (Cf. cependant [DIL 94]). Notre hypothèse de travail est qu'un modèle de recherche-utilisation de l'information est nécessaire, et peut être en partie dérivé des études empiriques réalisées jusqu'à présent.

Il faut pour cela bien définir ce qu'est l'analyse "cognitive" des tâches, et la distinguer d'autres approches possibles du problème. C'est ce que nous proposons de faire dans la partie qui suit. Puis nous examinerons les données empiriques disponibles afin de dégager les principaux facteurs pouvant influencer l'activité du sujet. A partir de ces études, nous proposerons quelques éléments pour un modèle général de l'activité de recherche d'informations.

2. La recherche d'informations : approche formelle ou approche cognitive?

2.1. Introduction

L'expression de "recherche d'informations" recouvre une grande diversité de situations et d'activités. Il n'en existe à l'heure actuelle aucune typologie basée sur des critères cognitifs, mais seulement des catégorisations générales reposant le plus souvent sur des métaphores. Par exemple Bernstein [BER 93] définit trois grandes classes d'activités basées sur le traitement d'informations documentaires, qu'il nomme extraction (*mining*), transformation (*manufacturing*) et jardinage (*farming*). Bernstein remarque que chacune de ces activités nécessite des outils informatiques spécifiques. Par exemple, l'extraction d'informations pourrait bénéficier d'outils de type "recherche automatique" mais peut-être pas d'outils de type "hypertexte" (pour une présentation en français de cette typologie voir [NAN 95]).

La conception d'environnements informatiques permettant d'assister les tâches de recherche d'informations demande une analyse détaillée de ces dernières. Jusqu'à présent ce problème a été abordé selon deux approches très différentes: d'une part une approche formelle, d'autre part une approche centrée sur l'activité de l'utilisateur.

D'un point de vue formel, toute recherche d'informations a pour but de localiser les informations pertinentes (compte tenu d'un objectif initial) tout en écartant celles qui ne le sont pas. Ceci correspond aux critères de "rappel" et de "précision" telles qu'elles sont définies par [SAL 83]. A partir de ces critères, il est possible de développer des méthodes de recherche et d'implémenter ces méthodes sous formes de structures de données, d'algorithmes et de langages de requête. Les travaux dans ce domaine ont utilisé le terme "hypertexte" principalement pour désigner des méthodes originales de structuration des données [RIJ 92] : systèmes à base de logique floue [CRO 93], réseaux de neurones [LEL 92] ou réseaux de Petri [STO 91]. Dans certains cas, les liens hypertextes ne sont qu'un outil parmi d'autres [GIR 92]. L'objectif de ces travaux est d'aboutir à des systèmes permettant la plus grande précision et rapidité dans la recherche. La question de leur mise en oeuvre par l'utilisateur humain et des problèmes qui peuvent s'y rattacher en est très souvent absente. Une conséquence en est que les systèmes sont souvent sophistiqués mais inadaptés aux besoins des utilisateurs. D'ailleurs, la plupart de ces systèmes ne connaissent qu'une diffusion très limitée, quand ils n'en restent pas au stade du prototype de laboratoire.

L'attention des psychologues et des ergonomes s'est récemment portée sur la façon dont l'opérateur se représente et exécute réellement une tâche de recherche d'informations. Dans cette nouvelle perspective, la définition formelle (ou "tâche prescrite") n'est qu'un élément parmi d'autres, et certains facteurs cognitifs tels que l'expertise du sujet doivent être pris en compte. Cette "approche-utilisateur" permettrait selon certains auteurs d'apporter des réponses plus satisfaisantes aux

problèmes de conception des systèmes informatiques (Cf. [AND 90] [MAH 93] [WRI 93]).

Wright définit plusieurs catégories d'activités de recherche d'informations basées sur les caractéristiques de l'information recherchée [WRI 90]. Cette information peut être simple et connue à l'avance, ou au contraire complexe et définie de façon progressive, à mesure de l'activité de recherche. De plus, Wright souligne la diversité des activités qui peuvent prendre place durant la recherche d'informations : lecture, prise de notes, évaluation et comparaison d'informations, etc. Chacune de ces composantes peut faire appel à des outils spécifiques, tels que système de prise de notes, multi-fenêtrage, références croisées ou marquage des sélections.

D'autres auteurs ont identifié certains facteurs situationnels et individuels qui peuvent rendre la tâche plus ou moins facile : expertise dans le domaine, capacités de lecture, familiarité avec le système [KNI 90] [ROU 92]. Nous reviendrons en détail sur ces facteurs dans la troisième partie de cet article. Auparavant nous proposons une synthèse des différentes approches de la notion de tâche de recherche d'informations.

2.2. Trois niveaux d'analyse des tâches de recherche d'informations

A la lumière des travaux évoqués ci-dessus, il apparaît que trois niveaux d'analyse des tâches doivent être distingués : d'une part, une description formelle de la tâche ; d'autre part, une représentation cognitive construite par le sujet ; enfin, une description de l'activité cognitive effectivement réalisée. A chaque niveau, il est nécessaire de définir les buts, les moyens et l'environnement nécessaires à la tâche. Le Tableau 1 présente la catégorisation ainsi établie.

	But	Moyens	Environnement
Modèle rationnel	Définition et structure formelles de but	Méthode et procédures optimales	Base de données et interface
Représentation cognitive	Représentation cognitive du but	Procédures et stratégies cognitives	Représentation du système d'informations (domaine, interface)
Activité cognitive	Régulation de l'activité	Exécution des actions	Information disponible

Tableau 1: Trois niveaux d'analyse des tâches de recherche d'informations.

Un **modèle rationnel de la tâche** décrit la façon la plus efficace d'obtenir l'information souhaitée, dans un environnement donné. Le modèle rationnel peut être défini indépendamment des caractéristiques d'un système cognitif particulier (connaissances, motivations, capacités, stratégies). Dans le modèle rationnel, le but représente l'état final, c'est à dire l'information à rechercher. Les moyens se composent d'un plan optimal mettant en oeuvre une série d'opérations, à la façon d'un algorithme. Enfin, l'environnement se compose d'une base d'informations, et d'une interface permettant l'exécution des actions. Il peut exister plusieurs modèles rationnels pour une même tâche [SIM 91].

Un **modèle cognitif de la tâche** décrit la façon dont un individu particulier se représente une tâche d'utilisation de l'information. La représentation cognitive du but peut être déterminée par une consigne externe, par exemple une question ou un énoncé de problème. Mais elle est également influencée par des facteurs cognitifs tels que l'expertise du sujet et son interprétation des énoncés. De la même façon, la stratégie cognitive (ou "plan") est influencée par, mais non isomorphe à la méthode rationnelle. Elle peut aussi être influencée par la représentation cognitive de but, et surtout elle peut être modifiée dynamiquement en cours d'exécution. L'environnement quant à lui correspond au système d'information tel qu'il est connu par le sujet. Cette connaissance peut être limitée à celle de l'interface du système, et elle évolue en fonction de l'expérience du sujet.

Le modèle cognitif de la tâche joue un rôle dans l'organisation (planification, contrôle) de l'activité du sujet, et fournit également des critères pour évaluer les résultats obtenus [CHA 94].

L'**activité cognitive** résulte de l'application d'un modèle cognitif de tâche dans une situation particulière. A ce niveau, le but est représenté par un mécanisme de régulation. La stratégie cognitive détermine la séquence des opérations cognitives qui seront réalisées. L'environnement est défini comme l'ensemble des informations et des outils de navigation disponibles à un moment donné.

Les relations entre ces trois niveaux sont tout aussi importantes que les niveaux eux-mêmes. Ainsi les résultats intermédiaires de l'activité peuvent conduire à une révision de la stratégie de recherche, de façon opportuniste [HAY 79]. Lorsque l'on étudie une tâche particulière, il est donc nécessaire de définir les correspondances, mais aussi les écarts possibles entre les entités définies à chaque niveau.

S'il est relativement facile de caractériser une tâche de recherche d'informations d'un point de vue formel, il est plus délicat de l'examiner du point de vue de sa représentation cognitive et de l'activité qui en découle. Cette approche est cependant indispensable pour comprendre à quelles difficultés peut se heurter l'utilisateur, et en déduire les caractéristiques d'un système réellement utile. Il faut pour cela recourir à l'expérimentation sur le terrain, et examiner en détail les variables situationnelles et individuelles qui influencent la conduite du sujet. La

partie qui suit résume quelques travaux menés selon cette perspective dans le domaine des hypertextes.

3. Aspects cognitifs des tâches de recherche d'informations

3.1. Introduction

Tout d'abord, à quelles difficultés l'utilisateur d'un système hypertexte peut-il se trouver confronté? Dans son article désormais classique, Conklin [CON 87] signalait déjà deux problèmes potentiels: la "désorientation" et la "surcharge cognitive" (*cognitive overhead*). Bien que Conklin n'ait pas proposé d'analyse approfondie de ces problèmes, leur réalité a été attestée par de nombreuses expériences [ROU 92].

La notion de désorientation repose sur une analogie entre l'hypertexte et un "espace sémantique" (voir [DIL 93], pour une analyse critique de cette analogie). Par opposition au texte imprimé, qui serait "linéaire" (donc uni-, voire au mieux bi-dimensionnel), l'hypertexte serait multidimensionnel, un "hyper-espace" informatif en quelque sorte. Pour utiliser l'information ainsi représentée, l'utilisateur doit construire une carte mentale du système, comme on le fait lorsqu'on visite une ville ou un immeuble pour la première fois. Suivant cette analogie, la désorientation correspondrait au fait de ne pas savoir "où" l'on se trouve, où l'on va, et comment s'y rendre. Elle indiquerait l'absence ou l'inexactitude de la représentation cognitive de l'espace informatif construite par le sujet. Ce type de phénomène a été mis en évidence notamment dans des travaux de Foss, sur lesquelles nous reviendrons dans la partie 3.3 [FOS 89].

La surcharge cognitive peut être définie comme un excès de traitements à réaliser ou d'informations à retenir. Elle repose également sur une analogie classique, celle qui compare le système cognitif humain à un système de traitement de l'information. Cette analogie est particulièrement mise à profit dans certains modèles de la mémoire de travail (voir par ex. [BRI 85]). Selon ces modèles, toute activité cognitive représente un certain coût, ou charge cognitive, qui est fonction de la nature et du nombre des opérations à réaliser dans un laps de temps donné. Or la lecture d'un hypertexte exige des opérations cognitives spécifiques, qui n'existent pas ou sont largement automatisées dans le cas du texte imprimé. Le sujet doit lire et comprendre l'information contenue dans une unité ou "noeud" de l'hypertexte, identifier les liens présents dans cette unité, et sélectionner l'un de ces liens afin de progresser vers une autre unité. Ces opérations peuvent s'avérer coûteuses et entraîner des problèmes de surcharge plus ou moins importants, en fonction notamment de certains paramètres de l'interface [WRI 91]. La surcharge

cognitive peut entraîner l'oubli intempestif des zones de l'hypertexte précédemment visitées.

Plusieurs expériences ont montré que la désorientation et la surcharge cognitive dépendent des caractéristiques de l'utilisateur et du format de présentation des informations (organisation des pages dans l'hypertexte, notamment). Cependant, étant donné la grande diversité des situations étudiées dans ces expériences, il est difficile de dégager des conclusions générales. Il faut au contraire considérer au cas par cas les différents domaines de tâches pour lesquels l'hypertexte peut être utilisé. Dans les sections suivantes nous allons aborder deux de ces domaines : la recherche d'informations spécifiques et l'exploration d'un corpus de textes.

3.2. Recherche d'informations spécifiques

Dans la perspective qui nous occupe, la recherche d'informations spécifiques peut être définie par deux caractères : d'une part, le sujet dispose au départ d'un objectif spécifique, qu'il peut avoir reçu de l'extérieur (question, énoncé de problème) ou bien élaboré lui-même. La réalisation de cet objectif nécessite le traitement (identification, compréhension) d'une quantité variable d'information textuelle. D'autre part, le sujet dispose de plus d'informations qu'il n'en faut pour atteindre l'objectif, et va donc devoir localiser et extraire les informations pertinentes.

Bien que les mécanismes cognitifs de la recherche d'informations soient encore mal connus, on sait qu'il s'agit d'une activité cognitive complexe, qui peut entraîner de nombreuses difficultés y compris chez l'adulte [GUT 88]. Une question essentielle est de déterminer si le format de présentation (linéaire, hiérarchique, en réseau, etc.) peut influencer la facilité avec laquelle l'utilisateur identifie l'information recherchée. Les recherches réalisées jusqu'à présent indiquent qu'aucun format n'est supérieur aux autres dans l'absolu. L'efficacité relative de tel ou tel format est influencée par deux types de facteurs : les capacités et stratégies de recherche d'informations d'une part, et la représentation de la structure d'ensemble des informations, d'autre part.

3.2.1. Capacités et stratégies de recherche d'informations

Rechercher des informations dans une base de données complexe nécessite des capacités cognitives qui ne sont pas nécessairement bien maîtrisées par des utilisateurs inexpérimentés. Dans une étude de Weyer [WEY 82], on demandait à 16 élèves de lycée d'utiliser un système d'information électronique ("Dynamic Book") pour répondre à une série de questions dans le domaine de l'histoire. Weyer observait que le traitement des questions complexes (par ex., questions nécessitant une comparaison) s'avérait difficile pour les élèves. De plus, les élèves éprouvaient

certaines difficultés à utiliser les outils de recherche les plus sophistiqués (par ex., une table de références croisées interactive).

McKnight, Dillon et Richardson [KNI 90] ont formulé des conclusions similaires à l'issue d'une expérience où 16 adultes devaient répondre à une série de questions en utilisant un document de 40 "pages" présenté sous forme imprimée ou hypertexte. Dans la condition hypertexte les sujets passent une plus grande partie du temps à examiner les menus (au détriment des informations de contenu), et utilisent rarement les liens directs entre pages (lesquels constituent un "outil" spécifique à l'hypertexte). Ces résultats indiquent que les outils censés faciliter la recherche ne sont pas spontanément maîtrisés par les utilisateurs. McKnight et ses collaborateurs suggèrent également que le format de présentation peut interférer avec les exigences de la tâche. Par exemple, les utilisateurs pourraient utiliser davantage les liens directs si la tâche consistait à résumer le contenu du document.

D'autres observations indiquent que la performance des utilisateurs peut s'améliorer à mesure qu'ils se familiarisent avec un système. Gray et Shasha [GRA 89] ont demandé à des étudiants de répondre à une série de questions de sociologie en utilisant un extrait de manuel présenté sur papier ou sur ordinateur, avec ou sans liens hypertextes. Les sujets du groupe "papier" parviennent à traiter plus de questions que les groupes travaillant sur ordinateur, et sont plus rapides pour quatre des cinq questions. Cependant, dans les deux groupes sur ordinateur le temps de réponse diminue significativement d'une question à l'autre, ce qui suggère un effet d'entraînement. De plus, dans le groupe hypertexte, l'utilisation des liens varie selon les sujets et les questions. Ceci indique de possibles différences dans les modèles de tâche construits par les sujets (nature de l'information pertinente et meilleure façon de la localiser). On pourra également consulter à ce sujet [WAN 88] ou [MAR 88].

Dans une série d'études récentes [ROU 94], nous avons également observé une amélioration des stratégies de recherche d'information chez des utilisateurs d'hypertextes sans expérience initiale. Dans la première expérience, soixante élèves de 12 à 14 ans ont participé à quatre sessions d'utilisation d'un hypertexte simple. A chacune de ces sessions, chaque élève devait chercher des informations dans l'hypertexte pour répondre à quatre questions sur un sujet d'intérêt général. L'efficacité de la recherche, estimée par un rapport de la qualité des réponses sur le temps passé à chercher l'information, augmente d'une session à l'autre, en particulier pour les questions les plus complexes. De plus, les élèves différencient peu à peu leurs stratégies de recherche selon la nature de la question. Par exemple, ils passent plus de temps à examiner les choix possibles lorsque la question nécessite une inférence. Il semble donc que les élèves acquièrent une meilleure représentation des exigences de la tâche ainsi que des caractéristiques du système.

Dans la seconde expérience, trente-neuf élèves de 16 à 18 ans devaient utiliser un hypertexte à structure hiérarchique (c'est à dire comportant un niveau

d'informations générales et un niveaux d'informations spécifiques) pour répondre à des questions plus ou moins complexes. Les résultats montrent que le temps de recherche diminue de la première à la dernière question d'une série, ainsi que de la première à la seconde utilisation de l'hypertexte. De nouveau, on observe que la stratégie de sélection varie selon le type de questions : pour les questions simples, les sujets sélectionnent essentiellement des unités "terminales", porteuses d'informations détaillées. Pour les questions plus complexes, ils sélectionnent une plus grande proportion d'unités sur-ordonnées et/ou non directement reliées. Ceci peut être interprété comme une stratégie d'exploration destinée à prendre en compte la plus grande complexité de l'objectif.

Une amélioration des stratégies de recherche avec l'entraînement est également observée dans une expérience de Tombaugh, Lickorish et Wright [TOM 87]. Tombaugh et ses collaboratrices ont demandé à des adultes sans expérience en informatique de lire un texte sur écran à l'aide d'un système de fenêtre unique ou multiple. La tâche consistait d'abord à lire le texte en entier, puis à y chercher l'information permettant de répondre à une série de questions. Dans la première expérience, les sujets ne recevaient qu'un minimum de familiarisation, et aucun avantage des fenêtres multiples n'a été observé. Dans la seconde expérience, un entraînement plus substantiel a été réalisé, à propos de la manipulation de la souris et des fenêtres. On observe alors un avantage du système de fenêtres multiples en ce qui concerne la vitesse d'accès à l'information-cible.

Ces résultats indiquent qu'une bonne représentation de la tâche et une bonne stratégie de gestion sont nécessaires pour accomplir des activités de recherche d'informations dans les hypertextes. Se représenter la tâche, c'est savoir quelle information il faut chercher pour répondre à une question donnée. Par exemple, reconnaître si la question nécessite l'identification d'une seule donnée, ou la comparaison entre plusieurs informations. Posséder une bonne stratégie de gestion, c'est pouvoir mettre en relation la tâche et les outils disponibles dans l'environnement. Par exemple, déterminer à un instant donné de la recherche s'il est préférable de consulter l'index, de sélectionner un lien direct ou de revenir à l'écran précédent. En l'état actuel des recherches, on sait encore peu de choses sur la nature de ces compétences, leur généralité, la façon dont elles peuvent être acquises et améliorées. Mais on sait qu'elles sont sensibles au contexte, et notamment aux informations qui signalent la structure d'ensemble de l'hypertexte.

3.2.2. *Effets des marques de structure sur l'activité de recherche*

Plusieurs études ont démontré qu'une représentation d'ensemble de la structure d'un hypertexte (par ex. un index hiérarchisé) pouvait faciliter la navigation et la recherche d'informations. Simpson et McKnight [SIM 90] ont cherché à évaluer l'effet d'indices structuraux sur la perception d'un hypertexte sur les plantes d'appartement. Ils ont comparé l'effet de trois types d'indices : une table des

matières alphabétique ou hiérarchique, une marque rappelant la dernière fiche ouverte ; la mise en majuscule des fiches sur-ordonnées dans la table des matières. Vingt quatre adultes (familiarisés avec l'informatique) devaient parcourir l'hypertexte une première fois avant de l'utiliser pour répondre à dix questions.

La table des matières hiérarchique est utilisée plus fréquemment que la table alphabétique ; elle entraîne une diminution du nombre de cartes ouvertes en lecture initiale comme durant la recherche d'informations (ce qui est interprété comme une diminution des ouvertures accidentelles ou erronées). De plus, elle permet aux utilisateurs de se rappeler plus précisément de l'organisation de l'hypertexte. Le marquage des sélections antérieures diminue le nombre de fiches ouvertes durant la lecture initiale mais pas durant la recherche d'informations. Les auteurs en concluent que les indices utiles pour le texte imprimé le sont aussi pour les hypertextes, et que différents indices peuvent être plus ou moins utiles selon les exigences de la tâche.

Le marquage de la structure peut avoir une influence sur la représentation de contenu que construit le sujet. Edwards et Hardman [EDW 89] ont demandé à 27 étudiants de traiter une série de 20 questions factuelles sur la ville d'Edinburgh en utilisant un hypertexte de 50 fiches. Trois formats de présentation étaient comparés : hiérarchique (les fiches sont reliées par des mots-clés insérés), index (les titres des fiches sont listés par ordre alphabétique dans un menu), et mixte (l'accès peut se faire soit par l'index soit par les liens). Les résultats montrent que le temps de recherche diminue de la première à la dernière question, cette diminution étant plus forte dans le format hiérarchique que dans les deux autres conditions. Ceci suggère qu'une organisation hiérarchique permet un meilleur apprentissage incident de la structure de l'hypertexte. La présentation hiérarchique favorise d'ailleurs le rappel de la structure de l'hypertexte. Elle permettrait donc la construction par les sujets d'une "carte cognitive" plus stable en mémoire.

Comme déjà signalé plus haut, les utilisateurs débutants ne tirent pas immédiatement avantage des nouveaux outils de navigation mis à leur disposition. Ainsi, Mohageg [MOH 92] a demandé à 64 adultes d'utiliser un hypertexte contenant des données géographiques pour répondre à une série de questions. Les données sont présentées selon quatre formats possibles : linéaire, hiérarchique, réseau ou mixte (hiérarchie plus réseau). Les sujets en condition hiérarchique et mixte mettent moins de temps à localiser les réponses que les sujets en condition réseau. De plus, les sujets n'utilisent pas toutes les ressources du système. Par exemple, dans la condition hiérarchique, 44% des sujets omettent d'utiliser la touche "retour" pour remonter à la carte sur-ordonnée. Ils refont pas à pas le parcours inverse.

Enfin, les outils de navigation hypertextes ne sont parfois utiles que pour certaines tâches. Wright et Lickorish [WLI 90] ont comparé l'efficacité de deux systèmes de navigation (l'index et la pagination) utilisés pour deux hypertextes

différents. Pour un hypertexte sur les plantes d'appartements de type "encyclopédie", l'index s'avère plus facile à utiliser et obtient la préférence des sujets. Pour un hypertexte de type "tableau de données", utilisé pour traiter des questions complexes, le mode pagination semble plus efficace.

D'autres expériences vont dans le même sens. Ainsi, Egan et al. [EGA 89] trouvent que par rapport à une version papier du même document leur système hypertexte ("Superbook") diminue le temps de recherche pour des questions dont les termes correspondent explicitement à une section du document. En revanche, le système s'avère plus lent pour des questions qui nécessitent une inférence. Mynatt et al. [MYN 92] aboutissent à des conclusions similaires à l'issue d'une étude où les sujets devaient utiliser une encyclopédie électronique pour répondre à des questions à l'aide d'un système de recherche automatique de chaînes. Le système s'avère plus performant que son équivalent papier pour répondre à des questions explicites.

Il est important de remarquer que dans ces deux dernières études les auteurs ont pris soin de fournir des explications détaillées ainsi qu'un entraînement substantiel aux sujets avant que ceux-ci n'accomplissent les tâches critiques. L'entraînement permet sans doute aux sujets de planifier leur stratégie de recherche sur la base d'une bonne représentation des potentialités du système.

Les résultats empiriques que nous venons de résumer confirment que la recherche d'informations spécifiques dans les systèmes hypertextes requiert des capacités cognitives adaptées. Tout d'abord, le sujet doit construire une représentation adéquate de la tâche. Par exemple, lorsqu'on lui pose une question, le sujet doit identifier la ou les informations permettant d'y répondre (un fait, une relation entre faits, etc.). Ensuite, le sujet doit savoir gérer la recherche de cette (ces) information(s). Les systèmes hypertextes pourraient assister la recherche d'informations en offrant des outils appropriés, tels que des cartes représentant la structure d'ensemble, ou bien des liens directs entre fiches. Cependant, l'exploitation de ces techniques nécessite que l'utilisateur adapte son savoir-faire, et qu'il développe des stratégies spécifiques. Les données empiriques montrent sans aucune ambiguïté qu'une bonne familiarisation avec le système joue un rôle important dans ce processus.

3.3. Exploration d'un corpus d'informations

Dans les recherches présentées ci-dessus, l'expression "recherche d'informations" a été utilisée pour désigner principalement des activités de localisation d'une ou de quelques informations, le plus souvent très simples. Cependant, dans bien des cas la "recherche" peut prendre des formes moins bien délimitées, notamment lorsque l'information-cible n'est pas précisément définie à

l'avance. L'objectif de l'utilisateur peut être d'explorer le système pour en dégager une connaissance globale (ce qui peut éventuellement lui permettre de mener des recherches plus précises dans un second temps).

La recherche d'informations peut également devenir difficile lorsque la complexité des buts ou la quantité d'informations à examiner augmentent. L'utilisateur (par exemple, l'élève qui explore un domaine de connaissances peu familier) doit alors planifier une **séquence d'étude** cohérente, ce qui pose parfois des problèmes avec les systèmes documentaires classiques [WAG 87]. Certains auteurs ont émis l'hypothèse que les hypertextes pourraient aider les élèves à organiser leurs propres séquences d'apprentissage, en facilitant la construction de parcours personnalisés [SPI 91]. Les différentes représentations ou structures d'accès (index, représentations graphiques, liens insérés...) permises par l'hypertexte permettent à l'utilisateur d'en explorer les portions intéressantes, ainsi que de construire une représentation cognitive de son organisation d'ensemble.

Il faut cependant souligner que, jusqu'à présent, cette hypothèse n'a pas reçu de confirmations empiriques définitives [ROU 92] [ROU inp]. Il existe d'ailleurs des obstacles théoriques à cette confirmation, comme Charney l'a fait judicieusement observer [CHA inp]. Par exemple, on sait que les séquences prédéfinies jouent un grand rôle dans la compréhension des textes imprimés. Le lecteur a tendance à considérer comme importantes les informations présentées en début de texte, et se base en partie sur la continuité référentielle pour construire une représentation cognitive intégrée du contenu d'un texte. On ne peut donc tenir pour acquis que l'exploration d'un ensemble d'informations sans structure préalable conduise à des apprentissages supérieurs à ceux obtenus dans des conditions comparables avec le texte imprimé.

Il y a donc deux problèmes essentiels à aborder : d'une part, il faut se demander dans quelles tâches d'exploration les hypertextes pourraient s'avérer utiles, compte tenu des capacités "ordinaires" des utilisateurs. D'autre part, il faut identifier les capacités spécifiques qui seraient à acquérir pour faire une utilisation optimale de ces systèmes.

3.3.1. Interaction entre contraintes de la tâche et formats de présentation

Tout comme les activités de recherche étudiées dans la section précédente, les activités d'exploration peuvent prendre des formes très différentes, selon que le sujet effectue un simple "survol" de l'information ou un examen approfondi, ou selon que la "région" utile de l'hypertexte est concentrée ou distribuée dans le système, etc. Il est important de déterminer précisément pour quel type d'exploration l'hypertexte peut être utile. Quelques études empiriques ont été consacrées à cette question.

Une expérience de Gordon et al. suggère que la présentation sous forme d'hypertexte n'est utile que lorsqu'un traitement approfondi de l'information est

requis [GOR 88]. Gordon et al. ont demandé à 24 étudiants de lire soit un texte d'intérêt général sans objectif d'apprentissage ("lecture d'agrément"), soit un texte technique avec un objectif de compréhension ("lecture finalisée"). Ces textes étaient présentés sur écran, selon un format linéaire ou hypertexte. La version hypertexte consistait à présenter d'abord une sorte de condensé du texte, où ne figuraient que les informations importantes. Les sujets pouvaient accéder à des détails supplémentaires en sélectionnant certains mots-clés du texte. La lecture du texte était suivie dans tous les cas d'un test de rappel et d'un questionnaire.

On observe que dans la situation de lecture d'agrément, la présentation linéaire entraîne un meilleur rappel des informations importantes du texte. Les sujets préférèrent d'ailleurs cette forme de présentation, et trouvent la version hypertexte plus difficile à utiliser. En situation de lecture finalisée, on n'observe aucune différence significative entre formats de présentation. Les auteurs concluent que l'hypertexte n'est sans doute pas adapté aux situations où la lecture ne vise pas explicitement l'apprentissage ou la mémorisation.

Un problème particulier révélé par cette étude est le sentiment d'incertitude généré par l'hypertexte. Certains sujets signalent en effet que face à l'hypertexte hiérarchisé, ils ont le sentiment de ne pas savoir "ce qu'il y a derrière la porte", c'est à dire à quelle information renvoient les mots-clés. On peut en déduire qu'une meilleure information sur la destination des liens contribuerait à faciliter la lecture de l'hypertexte.

Dee Lucas et Larkin [DEE inp] ont fait l'hypothèse que la présentation hypertexte peut faciliter la relecture d'un document, car elle autorise un accès direct aux sections que le sujet pourrait vouloir relire. Elles ont comparé trois formats de présentation d'un même texte sur l'électricité comportant neuf sections : format linéaire, hypertexte avec index alphabétique et hypertexte avec index hiérarchisé.

Dans la première expérience, 45 étudiants lisaient le document dans l'une des trois conditions avec pour objectif de se préparer à répondre à un test de compréhension. Les sujets devaient lire le document une première fois dans un ordre prédéterminé (le même dans les trois conditions), puis ils disposaient d'une période au cours de laquelle ils pouvaient relire certaines sections, s'ils le souhaitaient. Les auteurs ont constaté que les deux formats hypertexte entraînent un rappel du texte plus complet (les sujets rappellent des éléments issus d'un plus grand nombre de sections). De plus, la présentation structurée entraîne un meilleur rappel de la position des titres dans l'index.

Dans la seconde expérience, 63 étudiants devaient lire le même texte dans l'un des trois formats de présentation, mais cette fois avec un objectif plus spécifique : rédiger un résumé du texte. Les auteurs ont constaté que les sujets relisent plus de sections que dans la première expérience, et les différences entre formats de

présentation s'atténuent. Les auteurs en concluent qu'un objectif de lecture plus spécifique peut compenser des différences dans le format de présentation.

En résumé, la présentation d'informations sous forme d'hypertextes peut être plus ou moins efficace, en fonction de trois paramètres : la définition de la tâche de lecture, la représentation de la structure du document, et l'entraînement des sujets.

3.3.2. *Effets du format de présentation sur les stratégies de lecture*

Les études présentées jusqu'ici ont montré que les formats de présentation d'un même texte peuvent être plus ou moins favorables à l'activité de lecture. Mais la capacité des sujets à utiliser un format donné peut également dépendre de leurs compétences plus générales. Britt et al. ont observé une grande diversité de stratégies chez des étudiants utilisant un hypertexte simple pour la première fois [BRI in]. La tâche consistait à étudier un ensemble de documents relatifs à un problème historique. Certains étudiants suivent passivement l'ordre de présentation des documents dans l'index général, bien qu'ils aient été avertis que celui-ci est arbitraire. D'autres choisissent un ordre de lecture personnalisé mais non quelconque. En effet, dans ce dernier cas, on observe que certains types de documents sont sélectionnés en premier (par exemple, un extrait de manuel ou des documents primaires). De la même façon, seuls certains étudiants profitent du temps restant pour relire certains documents. Au total, tous ne tirent pas également parti des fonctionnalités de l'hypertexte.

Foss [FOS 89] a demandé à 10 adultes d'utiliser un hypertexte présentant des données géographiques pour effectuer une tâche impliquant la sélection et la mise en relation de plusieurs unités. L'hypertexte se composait de fiches présentant les caractéristiques de 10 pays selon 7 critères. Les fiches étaient "muettes", c'est à dire que le nom du pays concerné par chacune d'elles n'apparaissait pas. La tâche consistait précisément à identifier un maximum de pays dans un délai de 45 minutes, en lisant et comparant les informations contenues dans les fiches.

En examinant l'activité des sujets (ouverture/fermeture des fiches, disposition sur l'écran), Foss a pu identifier deux types de problèmes. Tout d'abord, certains sujets tendent à faire trop ou trop peu de comparaisons, ainsi qu'à oublier leurs hypothèses ou la façon dont ils sont arrivés à une conclusion. Foss attribue ce phénomène à une mauvaise "stratégie de recherche", c'est à dire une mauvaise représentation de l'objectif et des sous-buts et méthodes associés (par ex., procéder par élimination).

Ensuite, certains sujets tendent à ouvrir trop ou trop peu de fiches à la fois, ou bien encore à placer ces fiches sur l'écran d'une façon peu favorable aux comparaisons. Selon l'auteur, c'est alors une mauvaise "gestion" de l'activité qui est en cause, c'est à dire l'incapacité à mettre en oeuvre les moyens adéquats pour réaliser la tâche.

La représentation et la gestion de la tâche sont deux problèmes distincts, mais pas indépendants. Par exemple, une mauvaise gestion de l'activité (ouvrir trop peu d'unités) peut empêcher les sujets de mettre en oeuvre une représentation initiale de la tâche pourtant correcte. En résumé, cette expérience confirme que le succès d'une activité de recherche/utilisation d'informations dépend de la représentation des buts à atteindre, des moyens à mettre en oeuvre, et des caractéristiques de l'environnement dans lequel l'activité se déroule.

D'autres études ont montré un effet en retour, c'est à dire une adaptation des stratégies des sujets aux propriétés du format de présentation. [BLA 92] observent que le format de présentation de définitions "en-ligne", c'est à dire accessibles à partir d'un texte présenté sur écran, influence la fréquence avec laquelle les sujets sollicitent les définitions. Quand l'accès aux définitions est rendu plus facile (diminution du nombre de "clicks" nécessaires, maintien du texte à l'écran...) les sujets consultent plus volontiers ces dernières. Leur stratégie de compréhension s'ajuste donc en partie aux contraintes de l'environnement. De la même façon, [ROU 90] observe que le marquage des sélections antérieures réduit le phénomène de désorientation, comme le montre la diminution de la fréquence des "bouclages" (le fait de revenir accidentellement sur une unité précédemment visitée).

En résumé, l'adaptation des sujets à un format de présentation " non-linéaire " varie en fonction de leurs capacités cognitives générales. On observe cependant une certaine flexibilité des stratégies de prise d'information, qui peuvent s'ajuster aux contraintes ou aux opportunités offertes par l'environnement.

3.4. *Conclusions*

Nous avons établi que l'utilisation d'un hypertexte pose deux types de problèmes : d'une part, la construction d'une représentation de tâche adéquate, d'autre part la gestion de l'activité. Dans notre problématique, disposer d'une représentation de tâche adéquate signifie savoir comment obtenir le meilleur résultat en un minimum de temps et d'efforts, à partir des contraintes fixées initialement. Gérer l'activité signifie savoir comment tenir compte des propriétés et des contraintes de l'environnement, ici le système d'informations (taille de l'écran, outils de recherche, etc.).

Dans plusieurs cas, l'efficacité médiocre du format hypertexte peut être attribuée au manque entraînement des utilisateurs. A l'inverse, on montre qu'un entraînement permet aux utilisateurs de maîtriser les outils de navigation avancés [KNI 90] [MOH 92] [ROU 94]. Il est probable que l'entraînement autorise une meilleure adaptation de la stratégie de navigation en fonction de la représentation cognitive de la tâche. Une stratégie réellement efficace est en effet celle qui permet

d'exploiter les propriétés d'un environnement donné pour servir au mieux les objectifs fixés au départ (voir par ex. la notion de "environmental affordance" chez [WRI 93]).

Un résultat observé régulièrement est le besoin de représenter la structure d'ensemble du système. C'est par exemple la fonction des tables de matières hiérarchisées qui permettent de construire une meilleure vue d'ensemble du système, ce qui permettra ensuite de développer une meilleure stratégie de navigation. Cependant, les concepteurs doivent veiller à ne pas multiplier les représentations de haut niveau. Certaines études suggèrent en effet qu'il peut y avoir interférence entre plusieurs indices de structure [EDW 89] [TRI 90].

Afin de mieux sérier les problèmes, nous avons proposé une distinction entre deux modes d'utilisation de l'information : la recherche d'informations spécifiques d'une part, et l'exploration d'un corpus d'autre part. Si cette distinction est commode d'un point de vue méthodologique, il faudrait parler d'un continuum plutôt que d'une distinction nette entre ces deux domaines de tâches. Rechercher une information très générale ou distribuée dans un système revient à explorer une "région" de ce système. Ensuite, un même objectif de départ (acquisition de connaissances ou résolution de problèmes) peut inclure des sous-tâches d'exploration comme de recherche d'informations spécifiques. Des études récentes ont ainsi montré que les sujets peuvent accéder à une base d'information selon des modalités très variables au cours d'une même tâche [VRI 94].

Enfin, il faut noter que les études réalisées jusqu'ici ont utilisé des cadres théoriques et des méthodologies très divers. Pour cette raison, il n'est pas toujours facile ni même pertinent de comparer les résultats de différentes études. Cependant, ces études s'accordent sur un point : le système est d'autant plus efficace qu'il est adapté aux connaissances et stratégies que les utilisateurs possèdent et utilisent spontanément.

4. Lecture et recherche d'informations : évaluation, sélection et traitement

4.1. Introduction

Dans cette section nous allons examiner les relations entre représentation de la tâche et activité cognitive dans la consultation d'un hypertexte. Comme le montre la Figure 1, nous proposons de considérer la consultation d'un hypertexte comme un cycle de traitement constitué de trois phases principales : l'évaluation du but, la sélection et le traitement d'informations. Cette conception est compatible avec d'autres modèles de la recherche d'information [GUT 88] ou de l'interaction sujet-ordinateur [NOR 84].

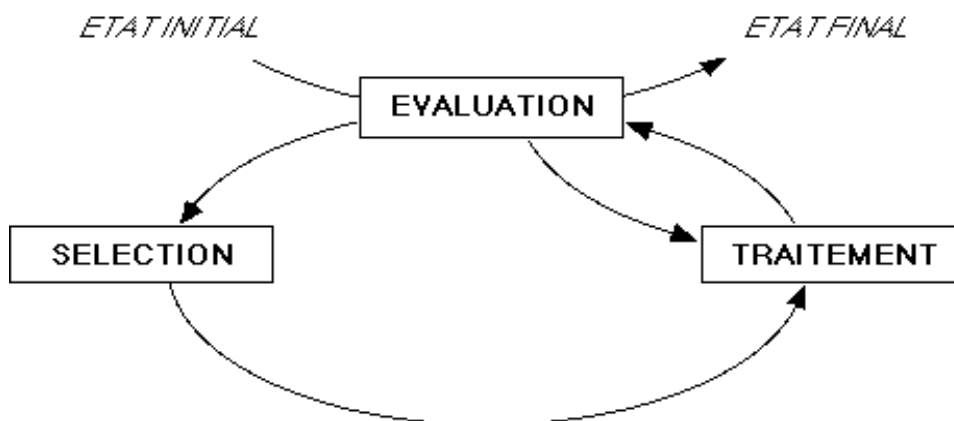


Figure 1: Le cycle Evaluation-Sélection-Traitement..

4.2. *Evaluation du but*

Comme nous l'avons suggéré plus haut, le modèle cognitif de la tâche inclut une représentation du but qui peut être modifiée dynamiquement. L'évaluation du but est le processus par lequel le sujet compare l'état actuel avec sa représentation du but. L'évaluation du but prend place au tout début de l'activité, quand le sujet définit son propre besoin d'information. Confronté à une question ou un énoncé de problème, le sujet doit d'abord décider s'il a besoin d'informations supplémentaires. Dans l'affirmative, l'évaluation le conduira à sélectionner une "catégorie" d'information pertinente (par exemple, ouvrir une base de données hypertextes et sélectionner un mot-clé dans l'index général).

Le maintien d'une représentation de but en mémoire peut interférer avec les autres composantes de la recherche d'informations (sélection, traitement), en raison de la capacité limitée de la mémoire de travail [BRI 85]. C'est en particulier le cas chez des utilisateurs en âge scolaire, qui ont souvent besoin de réactiver l'objectif défini initialement. [ROU 91] rapporte une série d'expériences dans lesquelles des élèves de 12 à 14 ans devaient utiliser un système hypertexte simple pour répondre à une série de questions. Dans une première version du système, la question était affichée en début de recherche, mais ne pouvait être relue ensuite. Cette particularité du système eut pour effet inattendu de rendre la tâche impossible pour un certain nombre de sujets. Ceux-ci lisaient en effet la question, mais l'oubliaient après deux ou trois sélections, et se trouvaient ainsi dans l'impossibilité

de continuer la recherche! Après que le système eut été modifié pour permettre la relecture de la question en cours de recherche, cette option fut sollicitée dans un tiers des cas environ. Ce pourcentage s'élevait même jusqu'à 40% pour les questions les plus complexes. Cet effet ne semble pas limité à la localisation d'informations précises.

Britt, Rouet et Perfetti ont observé des phénomènes similaires dans une population d'utilisateurs adultes [BRI inp]. Un groupe d'étudiants universitaires devait consulter les documents présentés sous forme d'hypertexte afin d'être de rédiger une courte synthèse à propos d'une controverse historique. La controverse était présentée sous la forme d'un énoncé qui apparaissait dans une fenêtre de type "pop-up", c'est à dire en surimpression par rapport au texte. Le sujet pouvait ainsi réexaminer l'énoncé à tout moment durant l'étude des documents. 68% des sujets en moyenne relisent l'énoncé au moins une fois. De plus, la relecture de l'énoncé n'intervient pas n'importe quand : elle se situe en début ou en fin de période dans 75% des cas. Une interprétation possible est que les sujets contrôlent ainsi leur activité, soit en planifiant leurs futures sélections, soit en vérifiant leur représentation du but après lecture. Ce résultat confirme que la représentation de but peut jouer à la fois un rôle d'organisation et d'évaluation de l'activité.

Une façon de réduire la charge liée au maintien du but est de garder une trace externe des résultats intermédiaires. Plusieurs techniques ont été proposées, et notamment des outils de prise de notes qui semblent faciles à utiliser, même par des utilisateurs débutants [OOS inp].

4.3. *Sélection d'un thème*

La sélection dans un hypertexte se fait par l'intermédiaire de menus ou de "boutons" insérés. L'utilisateur peut avoir à faire plusieurs sélections avant d'obtenir les informations désirées. Pour pouvoir faire ses sélections le sujet doit avoir une bonne représentation de l'information "cachée" derrière chaque option. L'ignorance de la destination des liens peut être un obstacle sérieux à la progression dans l'hypertexte, comme nous l'avons vu dans la partie 3.3.1. Il faut remarquer que le problème ne se pose pas dans les mêmes termes avec les documents sur papier. Même si le lecteur ignore où se trouve l'information, l'ensemble du texte est visible et peut être manipulé sans restrictions, ce qui n'est pas le cas avec les hypertextes. D'où la nécessité de fournir à l'utilisateur d'hypertextes une visibilité globale et des outils permettant de "survoler" l'hypertexte [LAI 91].

Dans quelques applications bien spécifiques, la destination des liens n'est pas un problème. C'est le cas par exemple dans un système expérimenté par Britt et al. [BRI inp]. Ce système permet la présentation d'un ensemble de documents relatifs

à un problème historique. L'utilisation des liens se limite à connecter les documents deux à deux par des références croisées. Par exemple, le sujet peut immédiatement consulter une source primaire citée dans un compte rendu. Le caractère univoque des liens élimine en grande partie l'incertitude quant à leur destination.

Dans tous les cas, la navigation nécessite un but, un plan pour atteindre ce but, et la capacité d'évaluer les résultats intermédiaires et de réviser le plan en conséquence si besoin. Une représentation correcte de la structure d'ensemble de l'hypertexte représente une aide appréciable pour l'orientation et la sélection.

4.4. *Traitement du contenu*

Toute sélection ou chaîne de sélections dans l'hypertexte aboutit *in fine* à la présentation d'un passage de texte dont le sujet doit lire et comprendre le contenu. Durant la phase de traitement du contenu le sujet acquiert une représentation de la signification du passage de texte sélectionné. On sait que comprendre un texte nécessite une hiérarchie de processus cognitifs. Au niveau local, le sujet doit comprendre la signification du texte, c'est à dire en extraire la macrostructure et l'intégrer à ses connaissances antérieures du domaine [DIJ 83]. Au niveau global, le lecteur doit évaluer si le passage lu contribue au but poursuivi, et le cas échéant intégrer l'information ainsi acquise à celles rencontrées précédemment (c.a.d. au résultat des cycles précédents. Voir [GUT 88]).

L'évaluation peut aussi avoir un effet en retour sur le traitement du contenu. Par exemple, le sujet peut décider d'interrompre la lecture si un certain seuil de pertinence n'est pas atteint par rapport aux objectifs. Enfin, il faut souligner que la transition entre les niveaux local et global de traitement doit être réalisée dans de bonnes conditions, faute de quoi le processus de compréhension peut se trouver perturbé [CHA inp] [ROU 90].

4.5. *Résumé et discussion*

Lors de la lecture d'un hypertexte, le traitement de base du texte est intégré dans un cycle complexe qui inclut des mécanismes d'évaluation, sélection et traitement du contenu. Bien que ces mécanismes s'appliquent également dans le cas des textes conventionnels, ils prennent un caractère obligatoire dans le cas des hypertextes. En effet, un hypertexte ne comporte par définition aucune organisation prédéfinie que le sujet pourrait suivre passivement. Au contraire, l'utilisateur doit construire lui-même une séquence cohérente. Cette nécessité d'un contrôle

stratégique peut expliquer en partie les difficultés observées chez des utilisateurs novices.

Le cycle d'évaluation, sélection et traitement n'a pas pour ambition de rendre compte de toutes les opérations cognitives qui prennent place lors de la lecture d'un hypertexte. Cependant, il fournit un cadre opératoire pour les notions de représentation de la tâche et de gestion de l'activité présentées dans la première partie de cet article.

5. Conclusion et perspectives

Dans cet article nous avons entrepris l'étude des activités de recherche d'informations sous l'angle de la représentation cognitive et de la gestion de ces activités. Nous sommes partis du constat selon lequel, jusqu'à présent, ces activités n'ont fait l'objet d'aucune approche théorique générale. Nous avons proposé une telle approche à partir des résultats expérimentaux obtenus jusqu'ici.

Nous avons montré qu'un modèle de tâche peut être défini à trois niveaux : la représentation formelle de la tâche, sa représentation cognitive et l'activité qui en découle. Les recherches présentées dans la partie 3 montrent que la représentation cognitive de la tâche et la gestion de l'activité sont des facteurs-clés, tant dans la recherche d'informations spécifiques que dans l'exploration d'un corpus. Nous avons proposé dans la partie 4 un modèle simple selon lequel la gestion de l'activité se décompose en trois phases : l'évaluation des buts, la sélection d'informations et le traitement de ces dernières. Bien que cette description ne soit qu'une proposition qui reste en grande partie à valider, elle permet d'expliquer en partie l'origine des difficultés que rencontrent parfois les utilisateurs d'hypertextes. Naviguer dans un hypertexte, c'est avant tout se construire des buts, les maintenir, et s'orienter dans le système afin d'y faire des sélections appropriées.

Nous pensons que l'analyse cognitive des tâches et activités de recherche d'informations peut être utile pour la conception de systèmes adaptés aux besoins des utilisateurs. Bien que nous soyons encore loin d'une méthode de conception rationnelle de ces systèmes, nous pouvons proposer quelques principes généraux, à partir des expériences réalisées jusqu'ici.

Tout d'abord, un système hypertexte doit permettre à l'utilisateur d'identifier l'organisation d'ensemble des informations. Cette organisation peut être représentée sous forme de "cartes" graphiques ou plus simplement de tables de matières structurées. Une représentation de haut niveau explicite permet à l'utilisateur de sélectionner les catégories d'information sur la base des objectifs poursuivis.

Ensuite, il est important de faciliter la progression dans l'hypertexte en évitant le phénomène de surcharge cognitive souvent observé. L'utilisateur devrait pouvoir

marquer les zones de l'hypertexte déjà parcourues, et sélectionner l'information pertinente à mesure où elle est identifiée. Les systèmes de prise de notes intégrés semblent particulièrement adaptés à cette fonction.

Enfin, il ne faut pas sous-estimer le besoin de formation qu'exige la prise en mains de tout système tant soit peu complexe. Les utilisateurs ont besoin d'apprendre les modes d'organisation de l'information, et le fonctionnement des outils permettant d'y accéder. Il leur faut ensuite adapter leurs stratégies de recherche/compréhension en fonction des propriétés nouvelles de ces outils.

En dépit des nombreux problèmes ergonomiques et cognitifs recensés dans la littérature, nous pensons que les systèmes hypertextes ont un potentiel en tant qu'outils pour l'éducation et la formation. Mais cet avenir passe par la prise en compte accrue des mécanismes cognitifs du traitement de l'information. Les réflexions présentées dans cet article se veulent un premier pas dans cette direction.

Bibliographie

- [AND 90] Anderson J.R., *The adaptive character of thought*, Lawrence Erlbaum, 1990.
- [BAL 90] Balpe J.P., *Hyperdocuments, Hypertextes, Hypermédias*, Eyrolles, 1990.
- [BER 93] Bernstein M., "Enactment in information farming", *Hypertext'93 Proceedings*, ACM Press, 1993, p. 242-249.
- [BLA 92] Black A. et al., "Consulting on-line dictionary information while reading", *Hypermedia*, 4, n° 3, 1992, p. 145-169.
- [BRI 85] Britton B.K. et al., "Cognitive demands of processing expository text: A cognitive workbench model", in B.K. Britton & J.B. Black (Eds.), *Understanding Expository Text*, Lawrence Erlbaum, 1985.
- [BRI inp] Britt M.A. et al., "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*, Lawrence Erlbaum, in press.
- [CHA 94] Chatillon J.-F. & Baldy R., "Performance motrice et développement moteur, les liens au développement cognitif", *Enfance*, 1994, n° 2, p. 299-319.
- [CHA inp] Charney D., "The impact of hypertext on processes of reading and writing", in S. J. Hilligoss & C. L. Selfe (Eds.), *Literacy and computers*, MLA, in press.
- [CON 87] Conklin J., "Hypertext: an introduction and survey", *IEEE Computer*, 20, n°9, 1987, p. 17-41.
- [CRO 93] Croft W.B. & Turtle H.R., "Retrieval strategies for hypertext", *Information Processing & Management*, 29, n° 3, 1993, p. 313-324.
- [DEE inp] Dee-Lucas D. & Larkin J.H., "Text representation with traditional text and hypertext", *Cognition and Instruction*, in press.
- [DEN 90] Denhière G. & Baudet S., *Lecture, compréhension de texte et science cognitive*, PUF, 1990.

- [DIA 89] Diaper D., *Task analysis for Human Computer Interaction*, Ellis Horwood, 1989.
- [DIJ 83] van Dijk T.A. & Kintsch W., *Strategies of Discourse Comprehension*, Lawrence Erlbaum, 1983.
- [DIL 93] Dillon A. et al., "Space - the final chapter or Why physical representations are not semantic intentions", in C. Mc Knight, A. Dillon & J. Richardson (Eds.), *Hypertext. A psychological perspective*, Ellis Horwood, 1993.
- [DIL 94] Dillon A., *Designing usable electronic text : ergonomics aspects of human information usage*, Taylor & Francis, 1994.
- [EDW 89] Edwards D.M. & Hardman L., "Lost in hyperspace: cognitive mapping and navigation in a hypertext environment", in R. Mc Aleese (Ed.), *Hypertext: Theory into practice*, Intellect Ltd, 1989, p. 105-125.
- [EGA 89] Egan D.E. et al., "Formative design-evaluation of SuperBook", *ACM Transactions on Information Systems*, 7, n° 1, 1989, p. 30-57.
- [FOS 89] Foss C.L., "Detecting lost users: Empirical studies on browsing hypertext", *Rapport de recherche INRIA n° 972*, 1989.
- [GIR 92] Girill T.R. & Luk C.H., "Hierarchical search support for hypertext on-line documentation", *International Journal of Man-Machine Studies*, 36, p. 571-585.
- [GOR 88] Gordon S. et al., "The effect of hypertext on reader knowledge representation", *Proceedings of the 32nd Annual Meetings of the Human Factors Society* (1988), Human Factors Society.
- [GRA 89] Gray S.H. & Shasha D., "To link or not to link? Empirical guidance for the design of nonlinear text systems", *Behavior Research, Methods, Instruments and Computers*, 21, 1989, p. 326-333.
- [GUT 88] Guthrie J.T., "Locating information in documents: examination of a cognitive model", *Reading Research Quarterly*, 23, n° 2, p. 178-199.
- [HAM 93] Hammouche H., "De la modélisation des tâches à la spécification d'interfaces utilisateur", *Rapport de Recherche INRIA n° 1959*, 1993.
- [HAY 79] Hayes-Roth F. & Hayes-Roth B., "A cognitive model of planning", *Cognitive Science*, 3, 1979, p. 275-310.
- [HOC 87] Hoc J.M., "Psychologie cognitive de la planification", Presses Universitaires de Grenoble.
- [JOH 91] Johnson P., "User interaction: a framework to relate Tasks, Users, and Design", in H.J. Bullinger (Ed.), *HCI'91*, Elsevier, 1991.
- [KNI 88] McKnight C. et al., "The construction of hypertext documents and databases", *The Electronic Library*, 6, n°5, 1988, p. 338-342.
- [KNI 90] McKnight C. et al., "A comparison of linear and hypertext formats in information retrieval", in R. McAleese & C. Green (Eds.), *Hypertext: State of the Art*, Intellect Ltd, 1990, p. 10-19.
- [LAI 91] Lai P. & Manber U., "Flying through hypertext", *Hypertext'91 Proceedings*, ACM Press, 1991, p. 123-132.

- [LEL 92] Lelu A. & Francois C., "Hypertext paradigm in the field of information retrieval: a neural approach", in D. Lucarella, J. Nanard, M. Nanard & P. Paolini (Eds.), *ECHT'92, Proceedings of the 4th ACM Conference on Hypertext*, ACM Press, 1992, p. 112-121.
- [MAH 93] Mahling D.E. & Croft W.B., "Acquisition and support of goal-based tasks", *Knowledge Acquisition*, 5, 1993, p. 37-77.
- [MAR 88] Marchionini G. & Shneiderman B., "Finding fact versus browsing knowledge in hypertext systems", *IEEE Computer*, 20, 1988, p. 70-80.
- [MAR 91] Marshall C.C. et al., "Aquanet: a hypertext tool to hold your knowledge in place", *Hypertext'91 Proceedings*, ACM Press, 1991, p. 261-275.
- [MOH 92] Mohageg M.H., "The influence of hypertext linking structures on the efficiency of information retrieval", *Human Factors*, 34, n°3, 1992, p. 351-367.
- [MYN 92] Mynatt B.T. et al., "Hypertext or book: Which is better for answering questions?", *Proceedings of CHI'92*, ACM Press, 1992, p. 19-25.
- [NAN 91] Nanard J. & Nanard M., "Using structured types to incorporate knowledge in hypertext", *Hypertext'91 Proceedings*, ACM Press, 1991, p. 329-344.
- [NAN 95] Nanard M., "Les hypertextes: au delà des liens, la connaissance", *Sciences et Techniques Educatives*, 2, 1995, p. 31-59.
- [NIE 90] Nielsen J., *Hypertext and hypermedia*, Academic Press, 1990.
- [NOR 84] Norman D.A., "Stages and levels in man-machine interaction", *International Journal of Man-Machine Studies*, 21, n°4, 1984, p. 365-375.
- [OOS inp] van Oostendorp H., "Studying and annotating electronic text", in J.-F. Rouet, J.J., Levonen, A.P. Dillon & R.J. Spiro (Eds.), *Hypertext and Cognition*, Lawrence Erlbaum, in press.
- [PAS 92] de La Passardière B. & Baron G.-L. (Eds.), *Hypermédiats et Apprentissages*, Presses de l'INRP, 1992.
- [RIJ 92] van Rijsbergen K. & Agosti M., "Editorial - The context of information retrieval", *The Computer Journal*, 35, n° 3, 1992, p. 193.
- [ROU 90] Rouet J.-F., "Interactive text processing in inexperienced (hyper-) readers", in A. Rizk, N. Streitz & J. André (Eds.), *Hypertext: Concepts, systems and applications*, Cambridge University Press, 1990, p. 250-260.
- [ROU 91] Rouet J.-F., Compréhension de textes didactiques par des lecteurs inexpérimentés dans des situations d'interaction sujet-ordinateur, thèse de doctorat, Université de Poitiers, 1991.
- [ROU 92] Rouet J.-F., "Cognitive processing of hyperdocuments: when does nonlinearity help?", in D. Lucarella, J. Nanard, M. Nanard & P. Paolini (Eds.), *ECHT'92, Proceedings of the 4th ACM Conference on Hypertext*, ACM Press, 1992, p. 131-140.
- [ROU 94] Rouet J.-F., "Question answering and learning with hypertext", in R. Lewis & P. Mendelsohn (Eds.), *Proceedings of IFIP WG3.3. workshop: Lessons from learning*, North Holland, 1994, p. 39-52.

- [ROU inp] Rouet J.-F. & Levonen J.J., "Studying and learning with nonlinear documents: Empirical studies and their implications", in J.-F. Rouet, J.J., Levonen, A.P. Dillon & R.J. Spiro (Eds.), *Hypertext and Cognition*, Lawrence Erlbaum, in press.
- [SAL 83] Salton G. & Mc Gill M.J. *Introduction to modern information retrieval*, McGraw Hill, 1983.
- [SIM 90] Simpson A. & McKnight C., "Navigation in hypertext: Structural cues and mental maps", in R. McAleese & C. Green (Eds.), *Hypertext: State of the Art*, Intellect Ltd, 1990, p. 74-83.
- [SIM 91] Simon H.A., "Cognitive architectures and rational analysis: comment", in K. van Lehn (Ed.), *Architectures for intelligence*, Lawrence Erlbaum, 1991, p. 25-39.
- [SPI 91] Spiro R.J. et al., "Cognitive flexibility, constructivism and hypertext: Random access instruction for advanced knowledge acquisition in ill-structured domains", *Educational Technology*, 31, n°5, 1991, p. 24-33.
- [STO 91] Stotts P.D. & Furuta R., "Dynamic adaptation of hypertext structure", *Hypertext'91 Proceedings*, ACM Press, 1991, p. 219-232.
- [THO 89] Thompson R.H. & Croft W.B., "Support for browsing in an intelligent text retrieval system", *International Journal of Man-Machine Studies*, 30, 1989, p. 639-668.
- [TOM 87] Tombaugh J. et al., "Multi-window displays for readers of lengthy texts", *International Journal of Man-machine Studies*, 26, 1987, p. 597-615.
- [TRI 90] Tripp S.D. & Robby W., "Orientation and disorientation in a hypertext lexicon", *Journal of Computer-based Instruction*, 17, n°4, 1990, p. 120-124.
- [TRI 95] Tricot A., "Un point sur l'ergonomie des interfaces hypermédia", *Le Travail Humain*, 58, n°1, 1995, p. 17-45.
- [VRI 94] De Vries E., *Structuring information for design problem solving*, thèse de doctorat, Université d'Eindhoven, 1994.
- [WAG 87] Wagner R.K. & Sternberg R.J., "Executive control in reading comprehension", in B.K. Britton & S.M. Glynn (Eds.), *Executive Control Processes in Reading*, Lawrence Erlbaum, 1987, p. 1-22.
- [WAN 88] Wang, X.H. & Liebscher P., "Information seeking in hypertext. Effects of physical format and search strategy", *Proceedings of the ASIS Annual Meeting*, 25, 1988, p. 200-204.
- [WEY 82] Weyer S. A., "The design of a dynamic book for information search", *International Journal of Man-Machine Studies*, 17, 1982, p. 87-107.
- [WLI 90] Wright P. & Lickorish A., "An empirical comparison of two navigation systems for two hypertexts", in R. Mc Aleese & C. Green (Eds.), *Hypertext: State of the Art*, Intellect Ltd, 1990, p. 84-93.
- [WRI 90] Wright P., "Hypertext as an interface for learners: some human factors issues", in D.H. Jonassen & H. Mandl (Eds.), *Designing hypermedia for learning*, Springer Verlag, 1990, p. 169-184.
- [WRI 91] Wright P., "Cognitive overheads and prostheses: some issues in evaluating hypertexts", *Hypertext'91 Proceedings*, ACM Press, 1991, p. 1-12.

- [WRI 93] Wright P., "To jump or not to jump: Strategy selection while reading electronic texts", in C. McKnight, A. Dillon & J. Richardson (Eds.), *Hypertext: a Psychological Perspective*, Ellis Horwood, 1993, p. 137-152.

Jean-François Rouet est chargé de recherche au CNRS, et poursuit ses travaux dans le domaine de la psychologie du langage, de la communication et de l'apprentissage. Ses recherches concernent principalement les mécanismes de compréhension et d'utilisation de systèmes documentaires complexes, tels que textes ou hypertextes à usage éducatif. Après avoir obtenu un doctorat de l'Université de Poitiers (1991), il a séjourné durant plusieurs années au Learning Research and Development Center (Pittsburgh), puis à l'Institut National de Recherche en Informatique et Automatique (Grenoble), avant de prendre ses fonctions actuelles au Laboratoire Langage et Communication (URA CNRS 1607, Université de Poitiers). Il a publié plusieurs articles et chapitres sur l'utilisation des systèmes hypertextes, et il a récemment coordonné l'édition d'un ouvrage collectif "Hypertext and Cognition" (Lawrence Erlbaum Associates, Publisher).