A Thomas.

Voilà un spécimen, avant le tirage officiel.
Il restera au moins quelques traces de notre passage du 2e au 3e millénaire, sur papier.
Amitiés.

JK

BULLETIN D'INFORMATIQUE APPROFONDIE ET APPLICATIONS

SCIENCE DE L'INFORMATION

N° 52 - MARS 1999

COMITE SCIENTIFIQUE

Patrick Abellard
Françoise Adreit
Jalal Almhana
France Chappaz
M'hamed Charifi
Roger Cusin
Bernard Goossens
Patrick Isoardi
Robert Jacquier
Jean - Philippe Lehmann
Nadia Mesli
Patrick Sanchez
Rolland Stutzmann
André Tricot

DIRECTEUR

Jean - Michel Knippel

REDACTEUR EN CHEF

Edmond Bianco

REDACTEUR ADJOINT

Sami Hilala

REDACTION

Université de Provence
Equipe Hermès. Case 33
3, place Victor Hugo
F - 13331 Marseille Cedex 3
Téléphone: (0)4 91 10 62 30
Télécopie : (0)4 91 50 91 10

DEPOSITAIRE

Université de Provence
Bibliothèque Vniversitaire
3, place Victor Hugo
F - 13331 Marseille Cedex 3
Téléphone: (0)4 91 62 44 16
Télécopie : (0)4 91 95 75 57

1 **EDITORIAL.**
Education nationale,

par Edmond Bianco

3 **LES ENJEUX D'UN DIALOGUE HOMME-MACHINE SUR INTERNET: L'HYPERDIALOGUE,**

par José Rouillard

21 **FOREVA, TUTORIEL D'EVALUATION AUTOMATIQUE DES COMPETENCES EN SCIENCES,**

par France Chappaz

33 **VOUZZAVEDIBISAR.**
La chanson du Mammouth,

par Edmond Bianco

D'ici quelque temps le bulletin aura ses informations sur WWW:
http://newsup.univ-mrs.fr http://www.u-3mrs.fr

Publication trimestrielle, gratuite, de l'Université de Provence.

Edition 1999 ISSN 0291 - 5413

EDITORIAL,

Education nationale.

Gauche, droite, gauche, droite ... Ce n'est pas de l'instruction militaire, non plus comme chantait Piaf :« Tu me fais tourner la tête, mon manège à moi c'est toi ...», mais simplement l'alternance au pouvoir de deux étranges variétés politiques, en apparence aussi inepte l'une que l'autre, au moins en ce qui concerne la gestion du pays, pour le reste ...

Sortez un député de l'hémicycle, bien malin serait celui qui, s'il ne le connaît pas, pourrait affirmer, sans se tromper, ce type-là est "de gauche", ou bien il est "de droite". Même les discours, qui développent une démagogie galopante ne sont plus un critère de sélection. Il reste tout juste les invectives qu'on se jette à la figure et qui font partie de la règle du jeu. On observe toutefois que ces invectives se font à la fois discrètes et abstraites, sans plus jamais effleurer les cas de dysfonctionnement que la "justice" a le plus grand mal à soulever.

Au fond, on a largement perdu de vue à quoi correspondent politiquement les concepts de droite et de gauche, qui ont une surprenante tendance à se retourner l'un vers l'autre, et à se mettre ainsi à tourner en rond, en dehors de tout contexte. Chantant comme une comptine: « C'est toi qui l'a fait ... non c'est pas moi, c'est toi ... »

Le monde, il est vrai est, au moins en apparence, devenu très complexe. Mais s'agit-il vraiment d'une complexité liée à l'évolution de la société ou plus simplement liée au jargon "technocratico-administratif"?. Ou encore liée au fait que voulant refuser d'admettre les évidences voire de les nier carrément, on se trouve projeté dans un univers complètement distordu où toute tentative de redresser quoi que ce soit ne fait que provoquer de nouvelles distorsions ailleurs.

De manière un peu schématique, je dirais volontiers que le concept politique de droite s'adapte bien à la manière dont est organisée notre société occidentale, à son idéologie de base, aux préjugés dont on a beaucoup de mal à sortir, construits sur les deux axiomes de base:

«Il n'est pas d'économie saine sans un profit maximum.»

«Il n'est pas de bonne société sans une économie saine.».

Au point qu'on se demande à quoi pensent une petite poignée de petits juges qui osent pourchasser de gros poissons pour des motifs aussi futiles que des abus et détournements de biens sociaux, délits d'initiés et autres broutilles.

Société du profit maximum, de l'"économie-rentabilité", de l'expansion économique, de la mondialisation du profit et du confinement individuel. Tout ce qui permet de mettre l'individu dans un état de "manque" est bon

pour le commerce. Symétriquement, il suffit que quelque chose se vende pour qu'on le fabrique. Mais cette assertion n'est pas aussi limitative qu'on pourrait l'imaginer, car si quelque inutilité se vend mal, on fait appel à l'artillerie lourde de la "Pub", autre marchandise qui se vend très bien.
Dans ce carrousel de mercantilisme, certains produits jouent un rôle particulier, c'est le cas des drogues, des armes en général et des mines antipersonnel en particulier, d'une certaine informatique, qui assujettissent, abrutissent, mutilent, sécrètent de la dépendance, de la corruption, bref sont une bénédiction pour le commerce, le vrai.

Que ferait dans cette mécanique bien huilée un citoyen trop critique, clairvoyant, soupçonneux. Or c'est précisément l'école qui est chargée de former le citoyen à la vie de citoyen, et il devient évident pour tous les Mammouths de la Terre que son rôle naturel est de façonner de bonnes pièces de machines afin de faire tourner cette belle industrie. D'où, d'immenses efforts pour rapprocher l'école du lieu qui en consomme la production.

Le concept de gauche s'oppose au concept de droite comme s'opposent les deux sens du mot "social". Ou social désigne les intérêts d'un tout petit groupe qui monte une affaire, ou bien il désigne les intérêts de tous les autres. On ne voit pas trop comment ces intérêts ne pourraient pas être opposés, c'est bien là que les grincements apparaissent. Tout naturellement la droite défend les premiers, alors que la gauche est censée défendre les autres. Quand le Président, de droite, déclare dans sa campagne électorale:
« Trop d'impôt tue l'impôt ...» et que, dès son arrivée au pouvoir il fait passer la TVA à 21%, tout est normal, c'est dans son rôle. Mais quand Jospin, martèle dans sa campagne qu'il faut arrêter de brader les entreprises publiques, et que, dès arrivé au pouvoir il fait mieux que son prédécesseur, et même que Madame Thatcher, alors on peut penser que la couleur dont il se pare n'est qu'une couche de mauvaise peinture qui s'écaille facilement.

Quand tout le monde est d'accord comme nos deux larrons en foire, quand la gauche mène une politique qui rend la droite complètement inutile au point qu'elle s'effrite et se dissout dans le néant, alors nous approchons du moment de la grande entente cordiale. Enfin sont écartés les gêneurs, les raisonneurs, les mécontents, les contestataires, c'est ici que le chômage fait montre de sa réelle utilité, on perçoit comme un doux ronronnement dans le meilleur des mondes. Les profits partent de la base et s'acheminent dans le sens d'une progression hiérarchique, pour s'accumuler vers les sommets à travers lesquels ils diffusent dans les nuées intercontinentales, pour retomber dans les lieux obligés de la stabilité et du profit bouillonnant, drogue, armes, organisations militaro-fascistes, et tous endroits à rentabilité et stabilité optimales. Le tout, sous l'œil vigilant des C.I.A., K.G.B. et autres destructeurs du "Rainbow Warriors".

Edmond Bianco

Les enjeux d'un dialogue oral Homme-Machine sur Internet
- L'Hyperdialogue -

José Rouillard

Laboratoire CLIPS-IMAG, Groupe GEOD
Campus Scientifique, BP 53
38041 Grenoble Cedex 9 - France

jose.rouillard@imag.fr

Résumé :

Internet et le World Wide Web sont maintenant des outils que nous utilisons quotidiennement. Le dialogue Homme-Machine (DHM) est aussi une discipline qui suscite un intérêt particulier auprès des chercheurs en informatique et en sciences cognitives. Dans cet article, nous présentons une réflexion autour de la rencontre entre le DHM et les nouvelles techniques de diffusion de l'information sur le réseau Internet. Nous dressons un rapide état de l'art du domaine et définissons une nouvelle approche du DHM dans les environnements hypertextuels, nommée l'hyperdialogue. Un premier modèle générique d'hyperdialogue est présenté, et de nouvelles perspectives de recherche sont envisagées, au vu des enjeux du DHM sur Internet.

Mots-clés :

Dialogue Homme-Machine, Multimodalité, World Wide Web, Hyperdialogue

Préambule

Communiquer avec autrui s'avère un mécanisme complexe, quelle que soit la situation et le contexte. Les chercheurs en traitement automatique de la langue naturelle (LN) étudient depuis longtemps ces phénomènes et tentent de les modéliser. L'idée de remplacer l'interlocuteur humain par une machine n'est pas nouvelle et depuis les années 60, de nombreux systèmes informatiques ont été développés dans cet objectif, avec plus ou moins de réussite. L'être humain communique grâce à ses capacités cognitives et sensori-motrices (vision, ouïe, toucher, odorat, goût, geste, parole). L'objectif de nos travaux est d'étudier les différents modes de communication entre l'homme et des dispositifs qui pourraient compléter (à défaut de remplacer) l'association actuellement prédominante clavier-souris-écran. Nous nous intéressons plus particulièrement à l'interaction verbale en LN pour la recherche d'informations électroniquement archivées sur le World Wide Web. Dans cet article nous présentons nos réflexions sur ce sujet, un rapide état de l'art du domaine et un concept particulier de DHM dans des environnements hypertextuels : l'hyperdialogue.

1. Introduction

Les quantités d'information et le nombre de serveurs sur le réseau Internet croissent de telle manière qu'il n'est plus possible de les traiter entièrement de façon manuelle. D'autre part, l'information est de plus en plus partagée entre de nombreux utilisateurs (novice, familier, expert, ...) qui doivent communiquer pour l'extraire ou pour la diffuser.

Ce constat nous amène à nous interroger sur la pertinence des interfaces actuellement disponibles pour manipuler des données informatiques. Les modes d'interactions traditionnels avec la machine (clavier, souris, écran) sont certes efficaces, mais nous pensons que d'autres interactions (comme l'interaction verbale souvent qualifiée de plus naturelle pour l'homme) pourraient être des alternatives aux modes de communication H/M déjà employés.

En recherche d'information, la pertinence d'un document est relative à la requête posée par l'utilisateur. Les Systèmes de Recherche d'Informations (SRI) sont majoritairement conçus au regard des capacités de la machine. Pour poser une "*bonne question*" à un système informatique, il faut au préalable comprendre comment il fonctionne et formuler une requête adaptée, sans quoi les réponses données ne seront pas toujours celles attendues par l'usager.

Pour résoudre ces problèmes, nous nous intéressons à des interfaces anthropocentrées, avec des modes d'entrée et de sortie de l'information complémentaires à ceux actuellement utilisés. L'être humain apprend à parler avant de savoir écrire. On imagine que ce moyen de communication serait plus efficace pour un DHM, pour autant que la machine *comprenne* ce qui lui est dit. Il s'agit alors d'étudier le DHM avec comme référent le modèle humain, de manière à adapter la machine aux capacités de l'homme.

Dans le domaine des SRI, les chercheurs travaillent sur différents modèles (booléen, vectoriel, hybride...) [SALTON et al. 83] pour la recherche de données archivées sur serveurs, et ils considèrent essentiellement les deux critères suivants :
- la précision : il ne faut présenter *que* l'information pertinente à l'utilisateur.
- le rappel : il faut présenter *toute* l'information pertinente à l'utilisateur.

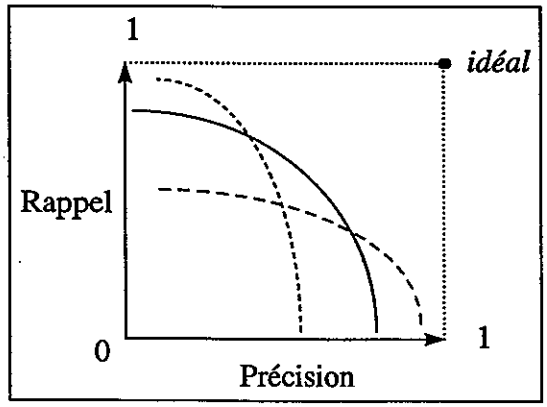


Figure 1 : Précision et rappel en recherche d'information : on souhaite obtenir simultanément *toute* et *que* l'information pertinente pour une requête donnée.

La **Figure 1** présente les notions de précision et de rappel en recherche d'information. Il est difficile d'obtenir la pleine satisfaction de ces deux critères simultanément. En ne voulant extraire que l'information pertinente, il se peut que l'on supprime de l'information qui aurait pu intéresser l'usager (on parle alors de *silence*), mais en s'efforçant de présenter toute l'information à l'utilisateur, on risque d'introduire de l'information non pertinente (on parle également de *bruit)* dans la réponse à la requête.

Bien avant les débuts d'Internet, les chercheurs étaient déjà confrontés aux problèmes de recherche d'informations optimales dans de grandes masses de données. On connaît désormais d'autres problèmes liés à la navigation dans des documents à structure hypertextuelle et hypermédia (comme le Web). Ces difficultés sont principalement la désorientation et la surcharge cognitive. Nous expliquerons de quelle manière le projet Orion, auquel nous participons, tente de répondre à ces problèmes.

David Sadek, à propos de nouvelles interfaces entre l'homme et la machine dit que *: "Toute offre de services doit intégrer comme préoccupation essentielle la convivialité de l'interaction usager/système, et ce, quels que soient les médias de communication, qu'ils soient traditionnels tels que la parole, l'écriture ou le geste, ou modernes tels que le clavier, la souris ou les outils de désignation tactile*" [SADEK 96].

Nous pensons qu'il est important de se poser des questions quant à la manière de concevoir et de réaliser les interfaces de demain en ce qui concerne l'accès aux informations, et particulièrement par le biais du réseau mondial Internet : Peut-on parler à une machine ? Dialoguer avec elle ? Pour quoi faire et comment le dire ? Quelle serait la pertinence d'une composante orale de dialogue pour Internet ? Les contraintes imposées par la mise en place d'interactions vocales peuvent-elles être compensées par une meilleure utilisabilité du Web pour l'usager ? En résumé, quels sont les moyens et les enjeux d'un dialogue entre homme et machine sur le réseau Internet ?

Nous nous intéresserons à l'utilisabilité de la parole comme moyen d'interaction entre l'homme et la machine (partie 2), puis à l'apport de l'interaction multimodale pour Internet (partie 3), et plus précisément de l'interaction verbale pour le WWW (partie 4). Enfin nous présenterons un modèle générique d'hyperdialogue (partie 5) avant de mesurer les enjeux d'un dialogue oral Homme-Machine sur Internet (partie 6).

2. Utilisabilité de la parole

Même si dans nos civilisations, la transmission des connaissances est basée sur l'écrit, il n'en demeure par moins vrai que l'un des moyens le plus naturel et sans doute le plus utilisé par l'être humain pour communiquer est la parole.

Depuis les débuts de l'informatique, nous utilisons principalement des périphériques spécifiques d'entrées de données dans le système, qui demandent un effort d'adaptation de la part de l'utilisateur. Si le clavier et la souris demeurent les outils les plus répandus, ils restent malgré tout artificiels, et demandent un temps plus ou moins important d'apprentissage pour un utilisateur novice.

Les avantages d'une interaction orale avec la machine sont nombreux, et beaucoup de chercheurs dans le domaine du dialogue oral Homme-Machine (H/M) les ont exposés. Françoise Néel, par exemple, rappelle que si l'on prend comme référence le modèle humain, les avantages apparaissent au premier abord évidents : *le naturel* (pas d'apprentissage particulier pour ce mode courant de communication), *la rapidité* et *l'efficacité* (débit d'information plus élevé et complémentarité avec d'autres modes, mémorisation, possibilité de macro-commande en un seul énoncé), l'*extension du champ d'action* (permet d'avoir accès à un objet non visible à l'écran par exemple) [NEEL 96].

Mais ce type d'interaction présente également des inconvénients. Si tous les utilisateurs ne disposent pas sur leurs bureaux, de microphones et de haut-parleurs pour dialoguer avec leurs machines, c'est que les contraintes imposées par un tel moyen de communication sont encore considérables. Du point de vue du traitement acoustique, les systèmes doivent être robustes pour faire face aux multiples difficultés rencontrées en situation réelle d'utilisation : environnement bruité, variabilité du signal d'un locuteur à un autre, capacité intrinsèque des supports matériels, puissance de calcul nécessaire pour une *bonne* reconnaissance. Ces limitations sont en grande partie techniques, et l'on suppose qu'avec l'amélioration des technologies, elles seront plus ou moins résorbées dans quelque temps.

Nous faisons donc l'hypothèse que l'aspect technique, même s'il est très important, ne doit pas nous empêcher de travailler sur l'aspect cognitif du DHM. Autrement dit, il ne *suffit* pas de reconnaître les mots prononcés, il faut aussi les comprendre, et en déduire l'intention de l'utilisateur. Nous rejoignons ici [AUSTIN 62] et [SEARLE 69] dans la théorie des actes de langages, qui se fonde sur le concept que "*dire c'est faire*" . En effet, dans la phrase "Pouvez-vous fermer la porte" , le locuteur ne souhaite pas une réponse de bas niveau comme "oui" ou "non", mais désire réellement faire faire un acte à son interlocuteur.

Les interrogations relatives à l'utilisabilité de la parole en interaction homme-machine sont :

(a) **Peut-on parler à une machine ?** Oui, répondent les chercheurs : "*Malgré leur complexité apparente, les systèmes de reconnaissance et de compréhension de la parole donnent des résultats satisfaisants dans des conditions d'emploi assez large (...) Il existe de très nombreux systèmes commercialisés, ce qui devrait permettre de voir de plus en plus d'applications vocales dans un proche avenir (serveurs vocaux, interfaces vocales, messageries, etc.)*" [CAELEN 96a]. En 1995, Philips sortait son système SP6000, avec reconnaissance de 26 000 mots en parole continue, tandis que Dragon Systems proposait une version multilingue avec une capacité de 120 000 mots.

Jean Caelen dit que "*reconnaître la parole, c'est la comprendre*" [CAELEN 96b]. Les capacités cognitives qui pour l'être humain sont naturelles doivent, si l'on veut les appliquer à une machine, reposer sur un modèle à forte orientation hypothético-déductive et d'interpolation par rapport à l'environnement, au contexte de la situation. Faisons une expérience empruntée à André Bisseret, montrant comment l'homme intègre les données reçues et les connaissances déjà acquises pour se créer une représentation mentale structurée [BISSERET et al. 96]. Lisez les deux phrases suivantes en essayant de visualiser la situation : *Jean allait à l'école. Il était tracassé par le cours de maths.* Mentalement, vous venez de créer un schéma selon les informations données. A présent, l'énoncé est complète par : *Il craignait de ne pas pouvoir contrôler la classe.* La simple reconnaissance acoustique de ces trois phrases par une machine ne suffirait pas à lui donner la compréhension de la scène, qui pour nous paraît évidente : il ne s'agit probablement pas d'un élève mais d'un professeur. C'est parfois aussi la culture d'un pays qui peut faire défaut pour la bonne compréhension d'un énoncé. C'était le cas pour cette chanteuse canadienne qui se produisait pour la première fois en France, et qui, entendant le public clamer "une autre, une autre" sortit de la scène en pleurs, croyant qu'il réclamait une autre chanteuse. Il y a un grand pas entre le décodage acoustico-phonétique d'un énoncé et son interprétation sémantique par la machine.

(b) **Veut-on parler à une machine ?** Globalement, la réponse semble être "Oui, si cela apporte quelque chose de plus qu'une interaction classique au clavier et à la souris". Nous montrerons dans le paragraphe 4. quels sont les freins et les motivations que nous avons étudiés auprès d'usagers du Web en situation de recherche d'information.

(c) **Pour quoi faire ?** Des tâches complexes, que nous nous efforçons de décomposer en une suite de sous-tâches pour que l'ordinateur fasse ce que l'on souhaite, seraient exécutées en un seul énoncé [ZANELLO 97]. Si l'ordinateur est capable de reconnaître et de comprendre l'énoncé "*imprime ce document en triple exemplaire recto/verso*", on conçoit aisément que ce genre d'interaction verbale puisse être préféré par certains à la suite de manipulation à la souris produisant le même effet.
En outre, pour ce qui nous intéresse plus particulièrement, le DHM devrait permettre de mieux cerner les besoins des utilisateurs lors de recherches sur Internet, et de s'adapter en fonction de leurs compétences. Il faut aller plus loin que le traditionnel "action/réaction" du clic souris. Avant de soumettre une requête à un moteur de recherche, puis de la reprendre, de la soumettre à nouveau, ne vaut-il mieux pas dialoguer avec l'usager à propos de son but profond, afin d'affiner sa requête et l'aider à mieux la construire ?

(d) **Et quoi dire ?** L'homme s'exprime en LN, et il est donc nécessaire, dans l'optique d'un outil anthropocentré, que la machine puisse reconnaître tout ce qu'il dit sans pour autant qu'il faille faire des pauses entre chaque mot ou articuler démesurément chaque syllabe.
Le vocabulaire utilisable doit être le plus large possible, et comporter à la fois des termes courants et des termes spécialisés pour la tâche à exécuter. Par exemple, un dialogue finalisé dans le domaine du renseignement SNCF doit pouvoir intégrer que la phrase "Je veux un train pour Paris" signifie en fait "Je veux une place dans un wagon de train pour Paris", ou "Je veux un billet pour un train allant à Paris".
De même, le système doit montrer sa compréhension du discours, et déduire certaines informations (par exemple que la gare de départ est celle où l'on se trouve lors du dialogue, si c'est une borne interactive).

3. L'apport de l'interaction multimodale au World Wide Web

"Aujourd'hui, l'efficacité recherchée n'est plus celle de la technologie prise isolément mais celle du couple "machine-humain". Il ne s'agit plus seulement d'exploiter au mieux les ressources techniques, mais de développer une technologie qui soit conforme aux attributs de la communication humaine. Et l'être humain voit, entend, parle, gesticule, manipule."[NIGAY et al. 96].

L'homme compose naturellement avec ses cinq sens. Le priver d'un ou de plusieurs de ses sens lors de l'utilisation d'une IHM semble impliquer que l'on va ainsi restreindre son champ d'activité, ou de perception. Des nombreuses recherches en IHM montrent que pour une tâche à accomplir donnée, la multimodalité est souvent facteur de réduction d'échec [ZANELLO 97].

Marie-Luce Bourguet donne dans sa thèse [BOURGUET 92] une taxonomie des systèmes multimodaux les plus communément admis, à savoir : multimodalité exclusive, concurrente, alternée et composée (ou synergique). Ces quatre types de systèmes multimodaux se distinguent selon que l'utilisation du mode de communication est séquentielle ou parallèle d'une part, et que la circulation de l'information sur ces médias se fait de manière combinée ou indépendante, d'autre part (Cf. Figure 2).

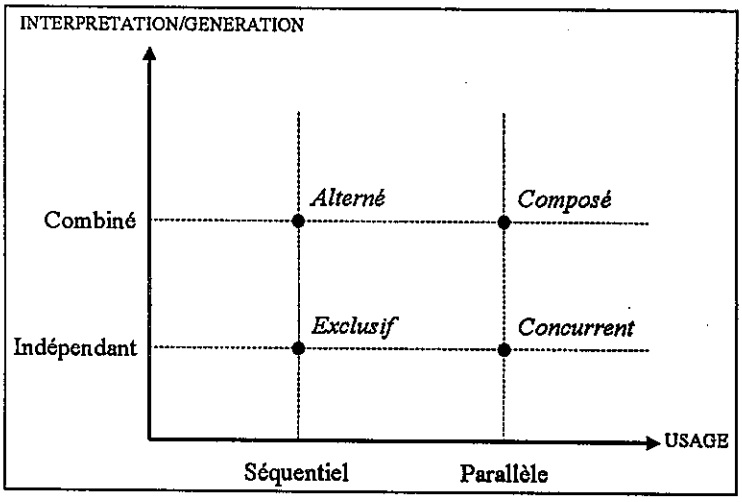


Figure 2 : Taxonomie des systèmes multimodaux : on distingue 4 types de systèmes multimodaux (exclusif, concurrent, alterné et composé) selon l'usage qui est fait des médias et selon les traitements effectués pour interpréter et générer les informations multimodales

En mode multimodal **exclusif**, deux médias ne peuvent pas être utilisés en même temps et les informations véhiculés par deux modes restent indépendants (exemple : clavier, puis souris). En mode multimodal **concurrent**, l'usage des médias peut se faire de façon parallèle, mais les informations circulant sur ces médias sont indépendantes. Il peut y avoir redondance d'information (exemple : affichage à l'écran et synthèse de la parole) ou conflit si deux commandes contradictoires surviennent en même temps. En mode multimodal **alterné**, l'usage des médias est séquentiel et le traitement des informations peut combiner différents modes (exemple : "mets ça là" [BOLT 80] en désignant l'objet et le lieu après la fin de la phrase).
En mode multimodal **composé ou synergique**, l'usage des médias s'effectue en parallèle et les traitements sont combinés (exemple : "mets ça là" en parlant et manipulant simultanément). Si ce dernier mode est le plus intéressant pour une communication naturelle, il est aussi le plus complexe à interpréter, notamment lors d'ambiguïté temporelles (dire avant de faire, faire avant de dire).

Les arguments théoriques en faveur de la multimodalité sont importants : on considère qu'une IHM multimodale doit permettre de maîtriser la surcharge cognitive impliquée par le nombre et la variété de données intervenant dans l'interaction. La communauté scientifique française a principalement travaillé sur la multimodalité dans les interfaces en entrée. L'objectif global de ces travaux est d'augmenter les capacités "sensorielles" des ordinateurs afin que l'utilisateur puisse communiquer en entrée plus naturellement avec la machine. D'autres travaux (notamment le Projet RICOM à l'IMAG) proposent d'exploiter la multimodalité également en sortie pour mieux rendre perceptible l'ensemble des informations manipulées par la machine. L'utilisation de plusieurs canaux de communication peut permettre d'attirer l'attention de l'usager sur un point précis si nécessaire. Par exemple, on pourra coupler la présentation d'une information de manière visuelle et auditive. La perception peut alors être

volontairement redondante ou bien complémentaire. Les avantages sensori-moteurs de chacune des modalités, doivent pouvoir compenser les inconvénients des autres prises séparément (comme le caractère volatil d'une information auditive, le besoin de précision d'un geste, etc.).

Nous avons vu plus haut quels pouvaient être les avantages et les inconvénients de l'utilisation de la parole (en entrée comme en sortie) dans des interfaces multimodales. Voyons à présent son utilisation dans le domaine de la navigation et de la recherche d'information sur Internet.

4. Vers une interaction verbale sur le World Wide Web

Le vaste monde d'Internet se trouve confronté à sa popularité grandissante. Dès la fin des années 80, des chercheurs mettent en évidence des problèmes rencontrés par les utilisateurs du Web. En effet, selon Conklin les deux plus importants problèmes liés à l'accès aux informations par le biais d'interfaces hypermédia sont la **désorientation** et la **surcharge cognitive** [CONKLIN 87]. Par ailleurs, Pitkow et Kehoe notent que ces difficultés ont principalement quatre origines possibles : "*The main problem people report when using the web are : (a) slow network or connection speed (b) not being able to find particular pages, even after they have been found before (c) not being able to manage or organize retrieved information and (d) not being able to visualize where they have been.*"[PITKOW et al. 96].

Hormis la vitesse d'accès (a), qui est du ressort des performances du matériel, les autres problèmes sont liés à l'interaction Homme-Machine (IHM). On voit ainsi se développer des travaux tendant à aider l'usager lorsqu'il se trouve confronté aux problèmes de surcharge cognitive et de désorientation sur le Web. [BOURDONCLE 97] propose ainsi une visualisation graphique des résultats obtenus lors d'une recherche d'informations sur le moteur Altavista[1] (Digital Equipment). De récents travaux tentent également de rendre le Web moins "impersonnel" notamment [EKLUND ET al. 96] et [BARRETT et al. 96] qui travaillent sur l'annotation (personnalisation) des pages visitées et les moyens d'échange de ces "méta-données" avec d'autres usagers du WWW.

Les travaux cités ci-dessus répondent aux problèmes de usagers en utilisant principalement un mode de communication écrit ou graphique. Il devient pertinent, voire fondamental, à présent, de prendre en compte la notion de multimodalité dans une IHM pour l'accès électronique aux documents. Nos recherches vont dans ce sens, et le projet Orion (Cf. 4.2) est actuellement au cœur de ces préoccupations. Ces recherches font suite à de précédents travaux dans le domaine de la navigation sur le Web grâce à la voix.

4.1. Naviguer sur le Web grâce à la voix

Dans le domaine de l'accès au WWW de manière vocale, des recherches ont été menées au *Center for Spoken Language Understanding* (CSLU) dans l'Oregon, aux États-Unis. En avril 95, David House, sous la direction du professeur David Novick, présente SLAM : *Spoken-Language Access to Multimedia*[2]. House décrit dans sa thèse une extension langagière à l'interface graphique du navigateur Mosaic. Son système SLAM (Spoken Language Access to Multimedia) est basé sur la constatation qu'interfaces orales et graphiques sont complémentaires l'une de l'autre :

"*Fortunately, many of the advantages and disadvantages of spoken-language-based-interfaces for hypermedia turn out to be complements of those for moused-based interfaces.*"[HOUSE 95]. Fort de cette remarque, David House propose une architecture dans laquelle un serveur distant permet la reconnaissance de la voix de l'utilisateur, et en réponse, fournit l'URL[3] adéquate à l'extension du navigateur Mosaic de chaque station de travail locale (client/serveur).

Mais toutes les pages HTML ne supportent pas SLAM car il faut qu'elles soient créées selon un protocole spécialisé (speech-capable document). L'inconvénient du modèle adopté par House réside également dans le fait que le serveur qui traite le signal est souvent très distant, et les accès sont lents (Portland, Oregon, USA). L'interaction entre l'utilisateur et la machine perd son caractère spontané et naturel.

Contrairement à SLAM, où il fallait appuyer sur un bouton pour pouvoir parler (processus touch to speak), Harouna Kabre et Jean Caelen avec Voiceweb [KABRE et al. 95] ont instauré un dispositif de reconnaissance vocale en tâche de fond (processus background Unix), qui est disponible à tout moment pour traiter du signal.

---

[1] http://altavista.digital.com/

[2] http://www.cse.ogi.edu/CSLU/publications/papers.html

[3] Uniform Resource Locator

C'est à dire que le logiciel détecte, selon des seuils déterminés, à quel moment l'usager commence et finit de parler dans le microphone.

La navigation grâce à la voix sur le Web ou dans des environnements hypertextuels est intéressante [HEMPHILL et al. 95], [HEMPHILL et al. 97], voire idéale pour certains [LAU et al. 97], mais la conclusion d'autres chercheurs qui proposent également un système de reconnaissance de la parole pour accéder au WWW montrent que "*le langage de requêtes est rigide et ressemble à un langage de commande par mot-clés*" [THIEBAUT et al. 96].

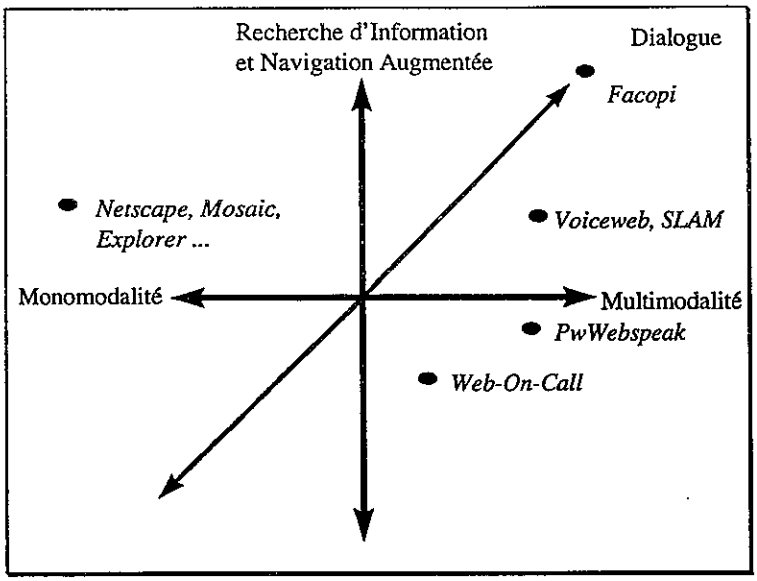


Figure 3 : Carte de positionnement des navigateurs

La Figure 3 présente une carte de positionnement des navigateurs actuellement disponibles pour la navigation sur le Web. On observe que les navigateurs du commerce (Netscape, Internet Explorer) ne permettent pas de piloter ou de dialoguer avec le système de manière multimodale. Notre équipe tente de se placer sur un point alliant une grande part de dialogue et de multimodalité. Nous avons proposé FACOPI (Feuilleteur A Commandes Orales Pour Internet). C'est un outil multimodal qui permet d'une part une navigation classique sur le web grâce à la voix (avant, arrière, etc.), mais aussi une recherche d'information au moyen d'un DHM [ROUILLARD et al. 97a]. Par exemple, à la question de l'utilisateur "je veux le cours de l'action Eurotunnel", la navigateur ne va pas lancer de recherche sur le Web sans avoir demandé à l'usager à quel domaine le terme "cours" fait référence (finance, pédagogie, etc.). Cela permet d'obtenir des réponses plus pertinentes puisque les requêtes sont mieux formulées.

Le système fonctionne par reconnaissance de mots-clefs, mais nous sommes encore loin d'un système de demande de renseignements en LN orale. C'est pourquoi nous orientons nos recherches, dans le cadre du projet ORION, vers les possibilités qu'offre le dialogue : reformulation, confirmation, sous-dialogues de négociation, etc.

4.2. Le Projet ORION

Le projet ORION est un programme de Recherche de l'ARASSH (Agence Rhône-Alpes pour les Sciences Sociales et Humaines - 1996). Il a pour objectif la conception de nouveaux composants logiciels de communication et d'interaction homme-machine susceptibles de faciliter la recherche d'information et la navigation dans le World Wide Web.

Cette recherche est motivée par le développement considérable que connaissent actuellement les réseaux et par le manque, dans les différents logiciels navigateurs commerciaux, d'aides graphiques, sonores et interactives avancées. Actuellement, il est difficile de retrouver de l'information efficacement sur le réseau Internet parce que les outils de navigation ne tiennent pas compte de la valeur sémantique des données manipulées. Par exemple, la requête "le cours de l'action Eurotunnel" soumise au moteur de recherche Altavista, donne un multitude des réponses non pertinentes du fait de la polysémie des termes de la requête.

Les difficultés de ce projet, qui se situe dans le champ de l'ergonomie de l'interaction personne-machine, sont beaucoup plus conceptuelles que techniques. La conception de cartes très spécifiques capables de représenter un espace informationnel hypermédia pour aider un usager non expert en informatique, ne peut être abordée avec succès que par une équipe pluridisciplinaire associant Sciences Humaines et Sciences pour l'Ingénieur. Les participants de ce projet apportent des compétences en Linguistique, Psychologie, Sociologie, Géographie, Sciences de l'éducation et Informatique.

L'une des voies que ce projet propose d'explorer est la conception de cartes cognitives multimodales pour le World Wide Web. Ces cartes ont pour vocation d'améliorer la perception par les usagers du WWW. Déjà quelques publications dans le domaine des cartes cognitives apparaissent par le biais du projet Orion [ZEILIGER et al. 97]. La notion de cartes cognitives nécessite à la fois : (a) une meilleure connaissance des comportements et processus cognitifs caractérisant l'activité de navigation pour une catégorie d'usagers bien déterminée et (b) une réflexion sur les modes de visualisation et d'interactions (vocal, gestuel, etc.) qui pourraient faciliter cette activité.

Les résultats attendus sont une meilleure connaissance des difficultés et comportements des usagers du Web en ce qui concerne la navigation et le repérage, les spécifications de cartes multimodales interactives pouvant faciliter l'accès du Web à un large public, en particulier dans le domaine de l'éducation dès le secondaire. Cela devrait nous conduire au développement de nouveaux navigateurs (également appelés "feuilleteurs" en français et "browsers" en anglais) répondant aux problèmes évoqués plus ci-dessus. Le pôle lyonnais (dirigé par Romain Zeiliger) s'occupe de la partie visuelle du problème, tandis que le pôle grenoblois (dirigé par Jean Caelen) travaille sur l'aspect oral et la mise en place d'un dialogue entre la machine et l'usager.

4.2.1. Projet Orion : l'aspect Visuel

Le pôle Lyonnais a choisi de travailler sur la partie visuelle des navigateurs, dans le domaine de l'éducation avec des enseignants et des étudiants. Depuis le début des années '90, dans la conception des logiciels, les approches centrées sur la technologie ont fait place - du moins sur le papier - à des approches centrées usagers. Un processus de conception commence généralement par une étude d'usages. Mais les usagers ne peuvent pas imaginer comment ils utiliseraient de nouvelles fonctionnalités technologiques parce qu'ils ne sont pas suffisamment familiers avec les innovations ou percées technologiques. Une étude des principes susceptibles de guider la conception d'aides à la navigation sur le Web a été menée [BELISLE et al. 97] , suivie de la réalisation d'un premier prototype logiciel : le navigateur NESTOR.

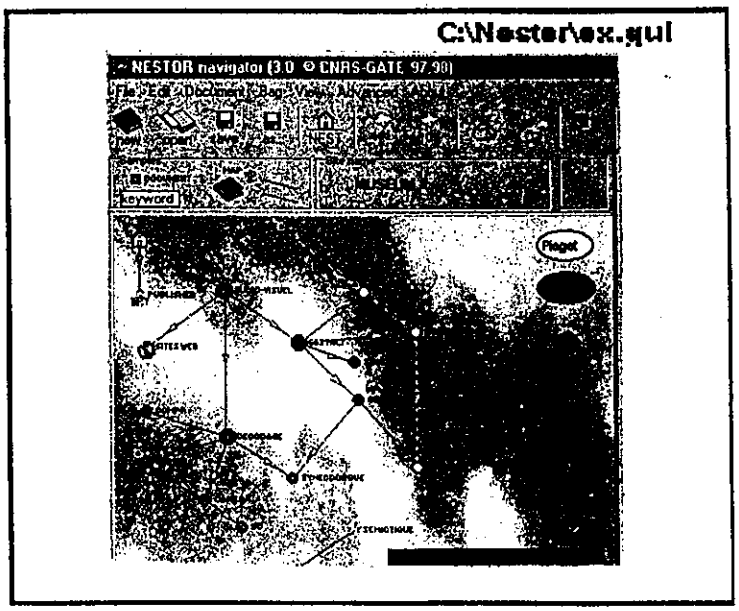


Figure 4 : Un exemple de carte de navigation du navigateur NESTOR

L'interface du navigateur NESTOR [ZEILIGER et al. 97] se présente d'abord sous la forme d'une fenêtre principale qui fonctionne comme celle des navigateurs standards. Trois fenêtres supplémentaires apportent chacune une aide spécifique : une fenêtre graphique "carte", une fenêtre "annotations personnelles", une fenêtre

"carte conceptuelle". Toutes ces fenêtres sont interactives et combinables : la "carte"- par exemple - se trace automatiquement dès que l'usager navigue avec une mise en page axée sur l'idée de "trajet". L'usager-apprenant peut re-disposer cette carte (par manipulation directe avec la souris) et l'utiliser pour retourner directement à un des sites déjà visités. Mieux, il peut graphiquement dessiner un lien nouveau entre deux documents quelconques : le lien hypertextuel obtenu - qui est personnel - apparaîtra automatiquement avec le document de départ et pourra être utilisé pour naviguer au même titre que les liens "publics" incorporés au texte original. Ces cartes peuvent être sauvées dans un fichier - comme les "bookmarks", retrouvées et partagées entre plusieurs utilisateurs.

Ainsi, NESTOR permet à l'utilisateur-apprenant 1) de visualiser graphiquement sa propre expérience de navigation en cours, et d'en reconfigurer la représentation, en fonction des objectifs poursuivis, 2) de progressivement passer d'une perception en termes de mots-clés à une catégorisation pouvant se représenter sous forme de carte conceptuelle, 3) de garder des traces de ces propres réactions mentales sous forme d'annotations textuelles liées. Cette facilitation de la navigation pour les usagers (novices notamment) est attribuée à l'aide apportée par la "carte" de NESTOR qui :

- grâce à son accès direct, diminue la charge cognitive dans les opérations d'exploration de type "centre et rayons"("hub and spoke") [CATLEDGE et al. 95].
- grâce à sa représentation contextuelle, diminue la désorientation dans les explorations de documents "éloignés" du point de démarrage (plus de trois niveaux "traversés").

4.2.2. Projet Orion : l'aspect Oral

Le pôle Grenoblois a choisi de travailler dans le domaine de la recherche documentaire; des enquêtes ont été réalisées auprès des usagers de la médiathèque de l'IMAG, de l'office du tourisme de Grenoble, et aux archives départementales de l'Isère.

Ces travaux doivent permettre l'intégration de l'usage de la parole et la mise en œuvre d'une composante de dialogue en sur-couche à la navigation sur le Word Wide Web. La majorité des interfaces homme-machine actuelles sont fondées sur une perception visuelle de l'information. Les navigateurs qui permettent d'évoluer sur le World Wide Web n'échappent pas à cette règle. Pourtant, nous savons que l'information disponible sur ce réseau est très volumineuse, et que l'affichage des réponses à une requête peut être un facteur de surcharge cognitive pour l'usager. Dans ce sens, des travaux montrent comment l'information peut être présentée de manière originale à l'utilisateur : Vernier et Nigay proposent une représentations multiples d'une grande quantité d'information avec leur logiciel *Vitesse*[4], basé sur une perception en œil de poisson (fisheye view) [VERNIER et al. 97].

Notre approche consiste, non pas à écarter ces canaux de communications, mais à proposer en sur-couche aux navigateurs actuels, des dispositifs audio d'entrée et de sortie de données, permettant une communication et un dialogue oral avec l'usager. Nous présentons tout d'abord les contraintes d'utilisabilité que l'on peut rencontrer pour une navigation en LN orale sur le Web, puis les résultats d'une étude relative à la motivation et aux attentes des usagers du web lors de recherches documentaires.

4.2.2.1. Contraintes d'utilisabilité

Afin d'étudier les relations entre les pages Web existantes et voir si rien ne s'oppose à la mise en place d'une architecture générique pour l'intégration d'une couche de DHM sur le réseau Internet, nous avons conçu et réalisé un robot logiciel (spider). Ce robot navigue automatiquement sur le web, en se propageant de liens (hypertextes) en liens, tout en sauvegardant à chaque nouvelle page visitée des données statistiques telles que le nombre d'ancres, le nombre d'images, l'origine de cette page, etc.

Dans un dialogue en LN, l'homme et la machine conversent idéalement dans une même langue. Or, nous savons que la réponse à la requête d'un usager peut se trouver n'importe où sur le réseau Internet. Plus globalement, on peut se demander jusqu'où peut mener une page HTML[5] si l'on suit tous les liens hypertextes qu'elle comporte et à quel moment va-t-on physiquement changer de pays (donc souvent de langue).

Grâce à cet outil, nous avons mené une étude sur la propagation au sein du Web à travers les liens hypertextuels [ROUILLARD et al. 97b]. Le robot informatique (spider) à été lancé sur le Web depuis six pages de base à

[4] http://iihm.imag.fr/demos/VitesseDemos.html
[5] HyperText Markup Language

travers le monde[6] de janvier à mars 1997. Les analyses des statistiques comparatives faites sur des pages issues de différents continents montrent la manière dont ces sites se font mutuellement référence. Il apparaît que tous les pays étudiés font référence à de très nombreux autres pays ou institutions comme on pouvait le supposer, mais de manière très disproportionnée d'un pays à l'autre.

L'étude montre également qu'environ 50% des images d'une page Web sont "cliquables". Cela signifie qu'elles ne sont plus seulement des éléments "décoratifs" pour capter l'attention de l'usager, mais qu'elles donnent, presque une fois sur deux, la possibilité d'interagir avec la machine. L'image prend de plus en plus d'importance sur le Web, et il faudra en tenir compte pour l'intégration d'une composante vocale. En effet, si le fait de passer d'un texte écrit à sa synthèse vocale ne pose pas de problèmes techniques particuliers, il n'en va de pas de même pour une image.

Enfin, nous avons mis en évidence que le DHM en LN doit intégrer des mécanismes de filtrage des données rapatriées depuis Internet pour éviter d'amener des ambiguïtés dans la communication. Nous avons démontré que l'on saute très facilement d'un pays à l'autre sur le Web : les données brutes (rapatriées en diverses langues) ne doivent donc pas interférer avec le processus de dialogue.

Les résultats obtenus jusqu'à présent dans le cadre du projet Orion nous indiquent que les utilisateurs ont besoin de logiciels de communication plus adaptés à leurs capacités cognitives. Nous avons observé des usagers du réseau Internet en situation de recherche documentaire. Nous donnons ci-dessous quelques résultats concernant cette étude.

4.2.2.2. Observation des comportements en recherche d'information documentaire

Hardie a montré qu'une grande majorité des recherches sur moteurs d'informations du Web se font à partir d'un terme unique pour la requête. Mais ce qui prédomine selon lui dans le comportement des utilisateurs d'Internet, c'est la faculté pour certains à élaborer une stratégie de recherche, en ayant à l'esprit le mode de fonctionnement des explorations faites par les robots de recherche. Les résultats de la recherche ne sont souvent fructueux que lorsque l'on connaît assez bien ce que l'on cherche. "*Even the most preliminary scans demonstrated that some users wanted the ocean, while others were looking for a grain of sand*" [HARDIE 96].

Dans le cadre du projet Orion, nous avons été amené à effectuer une enquête auprès des usagers de la médiathèque de l'IMAG (Institut d'Informatique et de Mathématiques Appliquées de Grenoble). Le but de l'enquête était d'étudier les possibilités de mettre en place une interface multimodale pour l'accès au catalogue de cette médiathèque via le Web. La figure 1 ci-dessous montre que pour ces utilisateurs, la recherche documentaire se fait principalement par tâtonnement au milieu des étagères et par le biais de bibliographies données par les professeurs : ils n'utilisent pas prioritairement l'accès au catalogue via le Web pour retrouver les documents dont ils ont besoin.

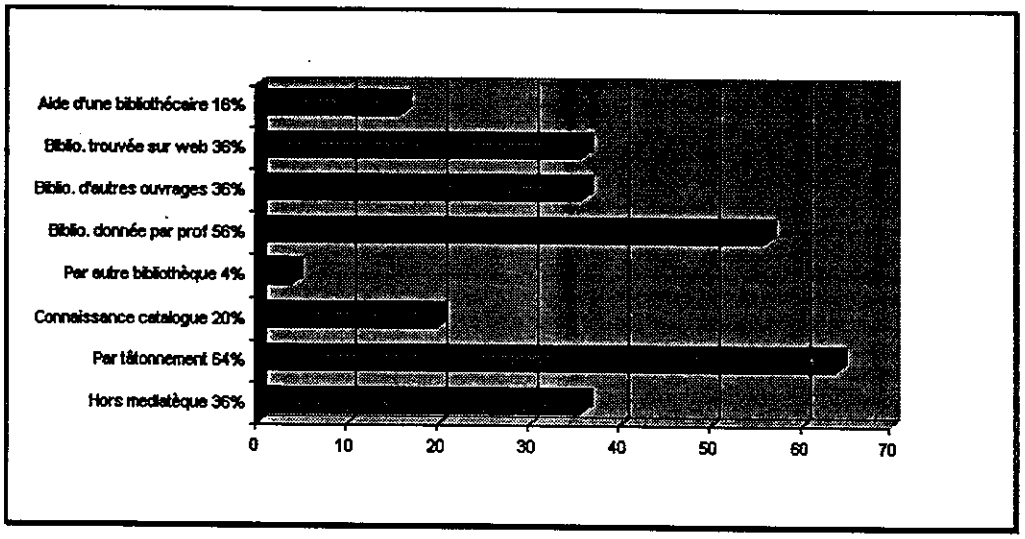


Figure 5 : Modes de recherche de documents des usagers (question à choix multiple)

[6] France, Allemagne, USA, Australie, Brésil, et Chine.

On observe que 68% des usagers n'utilisent les terminaux de recherche que rarement, voire jamais. L'enquête fait ressortir deux explications possibles :
- d'une part, des problèmes internes à la médiathèque : les terminaux ne sont pas en service, ne fonctionnent pas correctement, sont trop peu nombreux ou occupés (monopolisés) longtemps ;
- d'autre part, des réticences des usagers : ils n'apprécient pas l'interface proposée, ils préfèrent demander à un professeur une liste d'ouvrages, ils ne voient pas l'utilité de ce catalogue, etc.
Les utilisateurs sont favorables (68%) à une plus grande interactivité avec la machine grâce à un dialogue où ils pourraient dire par exemple : "non ce n'est pas cela que je recherche" ou "oui, je veux plus de détails", etc.
Une part importante des usagers se dit prête à tester une interface vocale pour une recherche documentaire (72%).

Même si les ouvrages sont présents et accessibles pour la plupart des usagers, ces derniers n'arrivent pas toujours à retrouver efficacement à partir de la page Web ce dont ils ont besoin. Ils dénoncent le moteur de recherche, pas assez précis ni pertinent à leur goût, et l'interface de ce catalogue, qu'ils trouvent peu conviviale. Les résultats de ces travaux sont disponibles sur Internet[7]. Pour apporter des solutions à ces problèmes de recherche d'information et de navigation sur le WWW, nous proposons une approche où le DHM rejoint la multimodalité dans le domaine des hypermédias. C'est le concept d'hyperdialogue (cf. partie 5). L'interaction verbale et le dialogue en LN peuvent, à notre avis, être des facteurs permettant de diminuer les risques de désorientation et de surcharge cognitive sur Internet.

4.3. Internet et interaction vocale

Pour certains chercheurs, l'interaction verbale est considérée comme une *modalité privilégiée* [JULIA et al. 97]. Nous présentons ici quelques recherches scientifiques et applications développées pour permettre une interaction vocale sur Internet et le World Wide Web. Nous décrivons les tâches et les types d'usagers susceptibles d'utiliser ces interfaces multimodales. Nous faisons un rapide rappel des logiciels déjà développés, et l'on discutera des contraintes d'utilisabilité que l'on pourrait rencontrer dans ce domaine.

4.3.1. Internet à la voix : pour quelles tâches ?

La messagerie électronique (E-mail) est l'un des outils les plus utilisés par les usagers d'Internet. De récents travaux [WALKER et al. 97] ont montré que l'utilisation de la voix pour piloter un logiciel de messagerie électronique est tout à fait réalisable. De plus, ces résultats confirment que la voix en commande d'entrée de dispositif favorise l'utilisabilité du système pour les personnes inexpérimentées dans la tâche à accomplir. Nous verrons plus loin que des travaux présentent également la multimodalité en sortie de dispositif. Concernant le Web, [ISSAR 97] explique comment on peut utiliser une interface vocale pour remplir un formulaire quelque soit le type de renseignements à fournir. D'autres travaux relatifs à la sécurité des réseaux utilisent la voix comme moyen quasi-infaillible afin de vérifier l'identité d'un usager du web [SOKOLOV 97].

4.3.2. Internet à la voix : pour quels usagers ?

En matière d'interaction multimodale pour le réseau Internet, le projet (MIWEB[8]) faisait état d'une population globale de 60 100 000 personnes ne pouvant pas utiliser une interface graphique traditionnelle, sur une population globale européenne de 800 000 000 d'individus (soit 7,51 %).

TYPE DE HANDICAP	EFFECTIF	%
Mains / Doigts	1 100 000	0,14
Avant - Bras	22 500 000	2,81
Coordination réduite	11 500 000	1,44
Dyslexie	25 000 000	3,13
Total	60 100 000	7,51
Population globale	800 000 000	100

Tableau 1 : Population cible du projet Miweb

"Ingram of Productivity Works estimates that in North America and Europe there are 27 million blind or visually people and that about 30 percent of them use computers." [9]

[7] http://hermes.imag.fr/~rouillar/mediatheque/enquete.html
[8] Multimodal Intercation with the WEb for the Blind
[9] *Webspeak' opens cyberspace to visually impaired* in The Times, by Mark Perkiss - 12.02.96

Une interface multimodale serait utilisée par la population citée dans cette étude, mais également par des usagers novices qui seraient probablement mieux guidés dans leur cheminement au sein d'Internet si le dialogue entre l'homme et la machine se rapproche de celui qu'ils connaissent et utilisent quotidiennement.
Il existe de très nombreux travaux à propos de la navigation sur Internet pour des personnes ne pouvant pas utiliser les outils communément employés pour cause de handicap moteur, visuel, etc., notamment [KENNEL et al. 96]. Les interfaces multimodales pour Internet sont appelées à se développer pour venir en aide aux personnes lésées par le manque d'accessibilité au WWW avec les interfaces graphiques actuelles [SMITH et al. 96].

En effet, quelques chercheurs comme [RAMSTEIN et al. 96] et [HERMSDORF 96] s'intéressent aux usagers non-voyants, malentendants, dyslexiques, et aux personnes ne pouvant pas utiliser leurs membres pour piloter une interface traditionnelle. Ils ont montré que l'intégration de la voix dans des navigateurs, comme moyen d'accéder à l'information disponible sur le Web, semble être un des concepts les plus prometteurs, puisque cela atténuerait en partie les problèmes de certaines personnes handicapées - principalement les aveugles [KRELL 96] - tout en permettant à terme une meilleure navigation (atténuation de la surcharge cognitive et de la désorientation), même pour les utilisateurs valides [PACIELLO et al. 96].

Il est aujourd'hui possible d'obtenir des informations du monde Internet, sans pour autant utiliser une interface graphique : Web-on-Call [RHIE 96] montre par exemple comment on peut "surfer sur le web" sans même posséder d'ordinateur. Le président de Netphonic explique :

"*Until now, a Web server could only be accessed by a computer with an Internet connection and Web browser software. Our new product makes Web content available to anyone ; with or without a computer. Using text-to-speech technology and audio recording feature called 'Teleprompting', Web-On-Call Voice Browser plays back Web documents over the telephone with a high degree of accuracy. It also supports documents retrieved via fax, e-mail and postal mail service.*"

Le projet pwWebSpeak, quant à lui, a été spécialement réalisé pour permettre aux handicapés visuels de naviguer sur le WWW sans avoir besoin de le voir : c'est un navigateur qui récupère les données et en calcule une synthèse vocale appropriée pour apporter un mode d'interaction entre l'homme et la machine ne se basant pas sur la vision et le graphique.

Pour aller plus loin encore, les chercheurs souhaitent concevoir et développer des systèmes capables de dialoguer avec un interlocuteur humain, de manière pertinente. Or, l'étude de la communication homme-machine se base principalement sur des observations de communication homme-homme et des expérimentations de type magicien d'Oz[10]. Pourtant, il existe bien des phénomènes différents qui interviennent dans une interaction homme-machine. Nous proposons une approche où le DHM rejoint la multimodalité dans le domaine des hypermédias. C'est le concept d'hyperdialogue. Il devrait être possible, à terme, avec l'hyperdialogue, de modéliser le DHM en s'appuyant sur des données réellement recueillies lors de dialogues entre l'ordinateur et l'être humain.

5. Un modèle générique d'hyperdialogue

A la jonction des hypertextes, du DHM et des sciences cognitives se trouve un domaine que nous proposons de nommer l'hyperdialogue (Cf. Figure 6). Nous définissons l'hyperdialogue comme étant un dialogue Homme-Machine, coopératif et finalisé, dans un environnement hypertextuel. Cet hyperdialogue peut être écrit, oral, voire gestuel. La multimodalité pouvant être effectuée avec usage des dispositifs en séquentiel ou en parallèle, et la circulation de l'information se faisant de manière combinée ou indépendante (Cf. [NIGAY ET al. 92] et [BOURGUET 92]).

Les sciences cognitives jouent un rôle essentiel dans l'hyperdialogue puisque nous nous positionnons dans une communication homme-machine anthropocentrée. Elles permettront de rendre compte des représentations mentales des utilisateurs et des différentes manières d'atteindre un objectif. L'hyperdialogue doit être finalisé, afin qu'un objectif commun anime les interlocuteurs et que l'on puisse juger à tout moment de l'état d'avancement vers cet objectif.

[10] Dans une expérience de type magicien d'Oz, un compère humain simule les comportements de la machine à l'insu des usagers. Ces derniers pensent donc dialoguer avec une machine, alors qu'en fait, il n'en est rien.

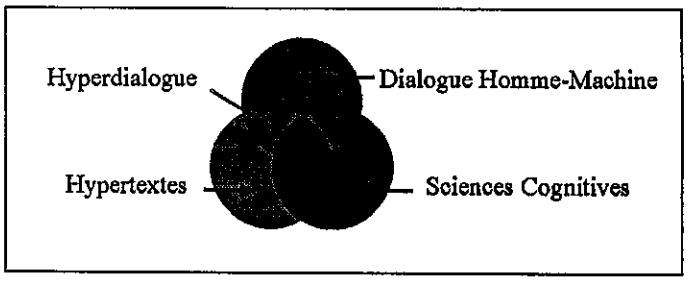


Figure 6 : Où se situe l'hyperdialogue ?

L'hyperdialogue doit être coopératif afin que les interactions soient optimales pour atteindre l'objectif fixé. Le principe de coopération des maximes de Grice [GRICE 75] préconise de *contribuer à la conversation de manière à correspondre aux attentes des autres interlocuteurs en fonction de l'état d'avancement et du but de la conversation.* Pour cela, quatre maximes sont retenues :

1) Maxime de modalité : être clair, éviter les expressions obscures, éviter l'ambiguïté, être bref et ordonné.
2) Maxime de pertinence : être pertinent.
3) Maxime de qualité : ne pas dire ce que l'on croit être faux, ne pas dire ce pour quoi l'on n'a pas de preuve suffisante.
4) Maxime de quantité : rendre la contribution aussi informative (mais pas plus) que nécessaire.

5.1. Principe de fonctionnement

Dans la Figure 7, nous présentons un modèle générique d'hyperdialogue dédié au WWW :

- Les éléments entrants peuvent être de plusieurs types : traditionnel (clavier+souris), verbal ou gestuel.
- Le module de reconnaissance verbale fournit en sortie une chaîne de caractères pour le module d'interprétation linguistique [KABRE 95a].
- Le processus d'interprétation gestuelle peut fournir des coordonnées et des mesures de déplacement.
- Les événements souris sont recueillis de manière traditionnelle (événement Windows par exemple).
- Une fusion des données est ensuite opérée avant l'analyse thématique des données. Le processus de fusion contrôle, à ce niveau, le type de fusion (redondance, complémentarité...) pour obtenir une analyse thématique cohérente.
- La procédure de dialogue s'engage, puis, une fois les ambiguïtés levées, le système peut envoyer une requête à un moteur de recherche sur le WWW. Sa réponse est alors analysée. Le cas échéant, des sous-dialogues de négociation, confirmation, etc. sont employés.
- Comme on avait opéré une fusion des données entrantes, nous proposons ici une fission des données sortantes. Ainsi, la présentation des résultats peut se faire de façon multimodale. Par exemple, une synthèse vocale peut prononcer les résultats en même temps qu'ils sont présentés de manière visuelle et graphique.

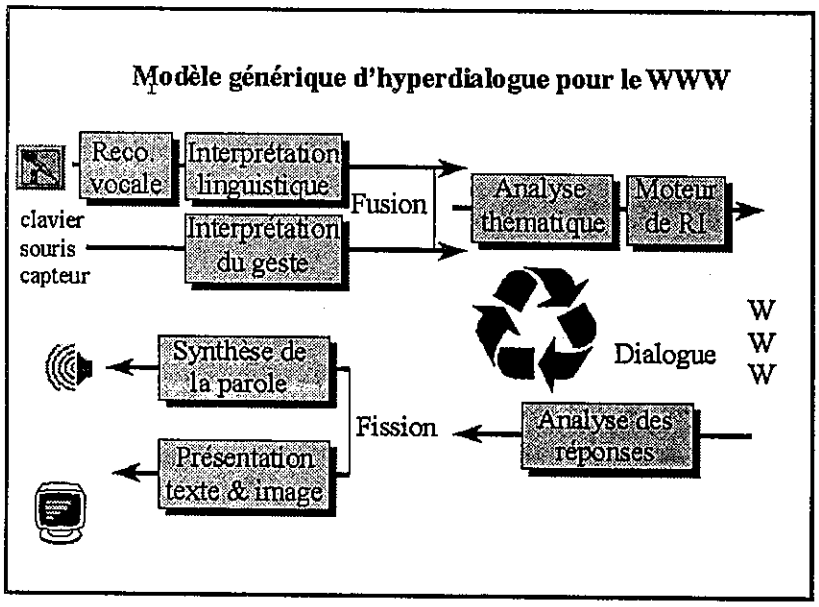


Figure 7 : Un modèle générique d'hyperdialogue pour le WWW

5.2. Interaction coopérative multimodale

Le cœur de ce modèle générique d'hyperdialogue réside dans l'interaction coopérative multimodale et finalisée que partagent la machine et l'utilisateur. Pour l'heure, une partie de notre équipe travaille sur la modalité particulière du canal auditif (reconnaissance et synthèse vocale), tandis que nous[11] nous intéressons à la composante dialogique de l'interaction.

Différentes stratégies de dialogues doivent être envisagées, selon les types d'usagers et les objectifs à atteindre. Une stratégie de dialogue est la manière de gérer les échanges entre interlocuteurs pour aboutir à un objectif. Cela consiste à maintenir un but, le déplacer, en proposer un nouveau, le différer, etc. Dans une stratégie **coopérative**, le dialogue va s'engager, les tours de parole s'enchaîner jusqu'à ce que la requête soit claire et précise. En stratégie **réactive**, la réponse est apportée immédiatement sans plus d'échange, quitte à faire des hypothèses. Il peut y avoir une stratégie **constructive** lorsque le sujet de la requête dévie du thème principal abordé. On rencontre également des stratégies **négociées** si les buts des interlocuteurs s'éloignent trop l'un de l'autre. Ils vont devoir négocier, chacun d'eux essayant d'obtenir une solution optimale.

Le dialogue est la clef de voûte de notre approche anthropocentrée, où la machine doit apprendre et comprendre les intentions de l'homme. La multimodalité, en entrée comme en sortie de données, est un moyen de permettre à l'homme de communiquer de manière naturelle avec la machine, en utilisant ses propres capacités sensori-moteurs et cognitives.

6. Conclusion

L'être humain compose plus naturellement avec ses cinq sens. Pourquoi le lui en ôter lorsqu'on lui propose d'interagir avec une machine ? Il faut adapter la machine à l'homme et pas l'inverse nous rappelle Jean Caelen en introduction des *Nouvelles interfaces Homme-Machine.* [CAELEN 96a].

Dans le domaine du dialogue oral Homme-Machine, les premières expériences (années 70) sont principalement basées sur la reconnaissance de mots isolés ou de mots clés. Depuis les années 80, le DHM tend vers des principes de coopération et une plus grande prise en compte du langage naturel. On cherche à interpréter les intentions de l'utilisateur. La théorie des actes de langage prend alors un rôle important pour la compréhension des énoncés. Les approches hypothético-déductives rendent le dialogue plus souple et moins artificiel. Dans un tel contexte, l'hyperdialogue va bénéficier à la fois des avancées technologiques en matière de télécommunications (débit des réseaux) et de celles des IHM. Finalement, l'enjeu d'un DHM sur Internet est double : économique et social.

6.1. L'Enjeu social

Nous avons vu que les personnes handicapées, (visuel, moteur : bras, mains, ...) ne sont pas exclues du monde d'Internet. Récemment, un appareillage de codification en braille a été mis au point par des chercheurs français, afin que des personnes aveugles puissent communiquer sur le réseau Internet. Le gouvernement français, conscient de l'ampleur du phénomène annonce des dispositions pour une diffusion plus large et plus uniforme des outils informatiques de demain[12]. Le DHM doit donc devenir un atout pour les communications sur Internet. La possibilité d'interagir avec la machine de manière multimodale doit être un avantage supplémentaire pour que l'homme interagisse en entrée et en sortie sur les données disponibles sur le réseau Internet.

6.2. L'Enjeu économique

L'enjeu d'un DHM sur Internet est de motiver l'usager en lui proposant un moyen d'interagir plus naturel (on n'a pas besoin d'apprendre à parler dans un microphone, contrairement à la manipulation du clavier et de la souris). Savoir présenter à l'homme toutes les facettes d'Internet (mail, WWW, accès aux documents grâce à la voix, etc.) suppose un niveau de dialogue important, mais les retombées au niveau industriel et économique risquent d'être considérables.

Le récent passage de l'état de monopole à celui d'oligopole qu'a connu le marché des télécommunications en France, témoigne de l'engouement des industriels pour les nouvelles technologies de l'information. Nous pensons que l'hyperdialogue sera l'un des phénomènes importants en matière d'IHM de demain. Plus l'on souhaite

[11] Jean Caelen, Solange Hollard et José Rouillard.

[12] 15 milliards de francs sur trois ans pour les nouvelles technologies de l'information et de la communication à l'école (Le monde Informatique n°743, 21 novembre 1997).

permettre l'accès aux nouvelles technologies au plus grand nombre et plus les systèmes proposés doivent être à même de répondre aux attentes de tous, experts mais aussi novices. Des travaux scientifiques balisent déjà ce terrain. On étudie comment un agent informatique peut tenir un discours cohérent en temps réel dans un système de messagerie électronique (chatting) [SUZUKI ET AL. 97].

6.3. Perspectives

Pour nous, le dialogue est vu comme un processus tendant à maximiser l'accord entre les interlocuteurs. Certes, mais que faut-il entendre par DHM en LN ? Ou du moins, que s'autorisent à dire les utilisateurs face à une machine censée comprendre leurs propos ? C'est ce que vise à déterminer le recueil du corpus HALPIN : Hyperdialogue avec un Agent en Langage Proche de l'Interaction Naturelle, que nous avons mis en œuvre au laboratoire CLIPS de Grenoble. Le système HALPIN fonctionne par reconnaissance de concepts dans le discours [ROUILLARD 98].

Cette étude devrait permettre d'accroître la pertinence des modèles actuels de DHM. On ne dispose en effet que d'un nombre restreint de corpus de DHM (et de faible taille). Une nouvelle méthodologie de recueil de corpus via le Web a permis d'obtenir rapidement un corpus original et varié. L'utilisateur est mis en situation de recherche d'information documentaire, et la machine est son partenaire pour cette tâche [ROUILLARD et al. 98]. Les perspectives de travaux pour améliorer le système consistent à lui intégrer un thesaurus afin d'obtenir une meilleure reconnaissance des concepts énoncés par l'usager d'une part, et l'utilisation d'un plus large vocabulaire lors de la génération de réplique, d'autre part. Nous avons déjà montré grâce à ce système qu'une partie du dialogue (les énoncés de la machine) peut être envoyée sur le réseau Internet pour que l'utilisateur entende son interlocuteur. Une étape importante demeure : nous devons mettre en place l'autre partie de cette communication orale, c'est-à-dire faire en sorte que l'usager puisse poser des questions de manière orale, sur le Web au serveur.

7. Remerciements

Nos remerciements vont à la **Région Rhône-Alpes**, qui nous a permis d'effectuer ce travail ; à madame **Françoise Renzetti**, et toute son équipe de la médiathèque de l'IMAG (Institut d'Informatique et de Mathématiques Appliquées de Grenoble) pour son accueil et ses conseils ; à mademoiselle **Dominique Lalane** pour ses multiples relectures et ses remarques toujours constructives.

8. Bibliographie

[ANTOINE 94] ANTOINE, J-Y., *Coopération synthaxe-sémantique pour la compréhension de la parole spontanée.* Thèse : INPG Grenoble, décembre 1994.

[AUSTIN 62] AUSTIN, J.L., *How to do things with words*, Oxford: Clarendon Press, Hermann. Traduit sous le titre *Quand dire c'est faire*, Seuil, 1970.

[BARRETT et al. 96] BARRETT, R., MAGLIO, P.P., KELLEM, D.C., *WBI: How to personalize the Web.* IBM Research Division, San Jose, 1996.

[BELISLE et al. 97] BÉLISLE, C., ZEILIGER, R., CERRATTO, T., *Integrated Cognitive Engineering at the Interface: A Tool Mediation Perspective* in Proceedings of the Second International Cognitive Technology Conference (CT'97), edited by J.P. Marsch, C.L. Nehaniv & B. Gorayska, Tokyo: IEEE Computer Society, 1997.

[BISSERET et al. 96] BISSERET, A., COUTAZ, J., NIGAY, L., *Assistance informatique aux archivistes départementaux. Analyses pour la future ergonomie du progiciel Ardent*, 1996. INRIA Rhône-Alpes.

[BOLT 80] BOLT, R.A., *Put-that-here : voice and gesture at the graphic interface.* Computer Graphics, 14 , 262-270, 1980.

[BOURDONCLE 97] BOURDONCLE, F., *LiveTopics : recherche visuelle d'information sur l'Internet*, RIAO'97 Computer-assisted information searching on Internet, Montreal, Quebec, Canada, 1997.

[BOURGUET 92] BOURGUET, M-L., *Conception et réalisation d'une interface de dialogue personne/machine.* Thèse, Grenoble, Septembre 1992.

[CAELEN 96a] CAELEN, J. *Reconnaître et comprendre la parole,* in Nouvelles interfaces Homme-Machine, Observatoire français des techniques avancées, OFTA, Paris, Décembre 1996.

[CAELEN 96b] CAELEN, J. *Reconnaître la parole c'est la comprendre* in Numéro spécial de la revue *La Recherche* n°285, "L'ordinateur au doigt et à l'œil" , p.62-65, Mars 1996.

[CATLEDGE et al. 95] CATLEDGE, L.D., PITKOW, J.E., *Characterizing browsing strategies in the World Wide Web, proceedings of the Third WWW conference,* Darmstadt, Germany, 1995.

[CONKLIN 87] CONKLIN, J., *Hypertext: an introduction and survey,* IEEE Computer, pp. 17-41, September 1987.

[EKLUND et al. 96] EKLUND, J., ZEILIGER, R., *Navigating the Web: Possibilities and Practicalities for Adaptive Navigational Support,* proceedings of AUSWEB'96, Gold Coast, Australia, 1996.

[GRICE 75] GRICE, H.P., *Logic and conversation,* in Cole, P., and Morgan, J.L., Syntax and Semantic, volume 3, Speech Acts. Academic Press, p. 41-58, 1975.

[HARDIE 96] HARDIE E., *A grain of sand or the ocean ; User aims in search engine interactions,* Fifth International WWW Conference - Poster Proceedings, INRIA / CNIT Paris La Défense - Mai 1996.

[HEMPHILL et al. 95] HEMPHILL, C.T., THRIFT, P. R., *Surfing the Web by voice, Proc. Of Multimedia '95,* San Francisco, November 1995.

[HEMPHILL et al. 97] HEMPHILL, C.T., YESHWANT K. M., *Developing web-based speech applications,* Eurospeech, Rhodes Greece, 1997.

[HERMSDORF 96] HERMSDORF, D., *Teleworking for disabled people - GMD's TEDIS project.* Fifth International WWW Conference (Workshop Web Accessibility for People with Disabilities) INRIA/CNIT, Paris La Défense, May 1996.

[HOUSE 95] HOUSE, D., *Spoken-Language Access to Multimedia (SLAM): A Multimodal Interface to the World-Wide Web,* Masters Thesis, Oregon Graduate Institute, Department of Computer Science & Engineering, April 1995.

[ISSAR 97] ISSAR, S., *A speech interface for forms on www,* Eurospeech, Rhodes Greece, 1997.

[JULIA et al. 97] JULIA, L. E., CHEYER, A. J., *Speech: a privileged modality,* Eurospeech, Rhodes Greece, 1997.

[KABRE et al. 95] KABRE, H., CAELEN, J., *VoiceWeb : une interface vocale pour l'accès au réseau Internet,* IHM'95, Grenoble, 1995.

[KABRE 95a] KABRE, H., *Echo User Guide,* CLIPS report no 1, 1995.

[KENNEL et al. 96] KENNEL, A., PERROCHON, L, DARVISHI, A., *WAB: World-Wide Web Access for Blind And Visually Impaired.* Computer Users. New Technologies in the Education of the Visually Handicapped, Paris, and ACM SIGCAPH Bulletin, June 1996.(http://www.inf.ethz.ch/department/IS/ea/blinds/)

[KRELL 96] KRELL, M. & CUBRANIC, D., *V-Lynx: bringing the World Wide Web to sight-impaired users.* The Second International ACM / SIGCAPH Conference on Assistive Technologies, ASSETS '96 - Vancouver, British Columbia, Canada, April 1996.

[LAU et al. 97] LAU,R., FLAMMIA, G., PAO, C., ZUE, V., *Webgalaxy – integrating spoken language and hypertext navigation,* Eurospeech, Rhodes Greece, 1997.

[NEEL 96] NEEL, F. *L'interaction vocale avec la machine : applications et perspectives,* in Nouvelles interfaces Homme-Machine, Observatoire français des techniques avancées, OFTA, Paris, Décembre 1996.

[NIGAY et al. 92] NIGAY, L., COUTAZ, J., *Interfaces multimodales et architectures : fusion et parallélisme.* IHM'92. Quatrièmes journées sur l'ingénierie des interfaces homme-machine, ENST, Paris, Décembre 1992.

[NIGAY et al. 96] NIGAY, L., COUTAZ, J. *Espaces conceptuels pour l'interaction multimédia et multimodale.* TSI, numéro spécial Multimédia et Collecticiel, Vol. 15, No 9,1996, pp. 1195-1225.

[PACIELLO et al. 96] PACIELLO, M., VANDERHEIDEN, G.C., LAUX, L.F., McNALLY, P.R., *Designing the World Wide Web for people with disabilities (Panel).* The Second International ACM / SIGCAPH Conference on Assistive Technologies. ASSETS '96 - Vancouver, British Columbia, Canada, April 1996.

[PITKOW et al. 96] PITKOW, J.E., and KEHOE, C.M. *Emerging trends in the WWW user population.* Communications of the ACM 39,6, pp. 106-108,1996.

[RAMSTEIN et al. 96] RAMSTEIN, C., MARTIAL, O., DUFRESNE, A., CARIGNAN, M., CHASSE, P., MABILLEAU, P., *Touching and hearing GUIs: Design issues for the PC-access system.* The Second International ACM / SIGCAPH Conference on Assistive Technologies. ASSETS '96 - Vancouver, British Columbia, Canada, April 1996.

[RHIE 96] RHIE, K., *Netphonic's Web-On-Call Voice Browser.* Fifth International WWW Conference (Workshop Web Accessibility for People with Disabilities). INRIA/CNIT, Paris La Défense, May 1996.

[ROUILLARD et al. 97a] ROUILLARD, J. et CAELEN, J. *A multimodal browser to navigate and search information on the Web.* Fourteenth International Conference on Speech Processing (ICSP97), IEEE Korea Council, IEEE Korea signal processing society. August 1997, Seoul, Korea.

[ROUILLARD et al. 97b] ROUILLARD, J. et CAELEN, J. *Étude de la propagation au sein du Web à travers les liens hypertextes.* Quatrième conférence Internationale Hypertextes & Hypermédias - Septembre 1997, Paris. Numéro spécial de la revue Hypertextes et Hypermédias, éditions Hermès, 1997, Paris.

[ROUILLARD 98] ROUILLARD, J., *Hyperdialogue Homme-Machine sur le World Wide Web : Le système HALPIN*, Colloque international Ergonomie et Informatique avancée ERGOIA'98, Biarritz, Novembre 98.

[ROUILLARD et al. 98] ROUILLARD, J. et CAELEN, J., *Etude du dialogue Homme-Machine en langue naturelle sur le Web pour une recherche documentaire*, Deuxième Colloque International sur l'Apprentissage Personne-Système, CAPS'98, Caen, Juillet 98.

[SADEK 96] SADEK, D., *Le dialogue homme-machine : de l'ergonomie des interfaces à l'agent intelligent dialoguant.* in Nouvelles interfaces Homme-Machine, Observatoire français des techniques avancées, OFTA, Paris, Décembre 1996.

[SALTON et al. 83] SALTON, G, McGILL, M.J., *Introduction to modern Information Retrieval,* McGraw-Hill, 1983.

[SEARLE 69] SEARLE, J.R., *Speech acts*, Presses universitaires de Cambridge. Traduit sous le titre *Acte de langage*, Hermann, 1972.

[SMITH et al. 96] SMITH, A., DUNAWAY, J., DEMASCO, P., PEISCHL, D., *Multimodal input for computer access and augmentive communication.* The Second International ACM / SIGCAPH Conference on Assistive Technologies. ASSETS '96, Vancouver, British Columbia, Canada, April 1996.

[SOKOLOV 97] SOKOLOV, M., *Speaker verification on the world wide web*, Eurospeech, Rhodes Greece, 1997.

[SUZUKI et al. 97] N.SUZUKI, S.INOKUCHI, K.ISHII AND M.OKADA, *Chatting with Interactive Agent,* Eurospeech, Rhodes Greece, 1997.

[THIEBAUT et al. 96] THIEBAUT, E., MARI, J.F., HATON, J.P., GONG, Y., FOHR, D. *Utilisation d'un système de reconnaissance de la parole pour accéder à W3.*, XXIes Journées d'Étude sur la Parole, Avignon, 1996.

[VERNIER et al. 97] VERNIER, F. et NIGAY, L., *Représentations Multiples d'une Grande Quantité d'Information*. IHM'97, Futuroscope de Poitiers, France, septembre 1997.

[WALKER et al. 97] WALKER, M., HINDLE, D., FROMER, J., DI FABBRIZIO, G., MESTEL, C., *Evaluating competing agent strategies for a voice email agent,* Eurospeech, Rhodes Greece, 1997.

[ZANELLO 97] ZANELLO, M.L., *L'utilisateur et l'interface multimodale. Contribution à la connaissance sur son utilisation et sa gestion,* Thèse sciences cognitives, Grenoble, juillet 1997.

[ZEILIGER et al. 97] ZEILIGER, R., REGGERS, T., BALDEWYNS, L., JANS, V., *Facilitating Web Navigation: Integrated tools for Active and Cooperative Learners*, ICCE'97 Conference, Kuching, Malaysia, December 97.

ForEva, Tutoriel d'évaluation automatique des compétences en sciences; réalisation et fondements

France CHAPPAZ
Maître de conférences

Equipe Hermès - Représentations et Transferts de connaissances
Université de Provence (Aix-Marseille I)
3, place Victor Hugo
13331 MARSEILLE Cedex 3 - France

fchappaz@newsup.univ-mrs.fr

Résumé :

L'auteur présente un tutoriel d'évaluation des connaissances et des compétences en Physique. Dans sa version ForEva en cours de réalisation, le tutoriel, à l'aide d'Internet, doit permettre l'élaboration à distance de bilans concernant l'activité scientifique liée à la résolution de problèmes au sens cognitif. Une version destinée au fonctionnement présentiel permet de mettre au point la méthodologie d'un traitement automatique des compétences.
L'auteur, à partir d'un modèle des opérations qui interviennent dans le processus de résolution, propose de substituer dans ce type d'activité mentale, la notion de fonction à celle de compétence : appliquée à un argument de connaissances, ce faire-cognitif, en transformant l'objet étudié, permet la performance.

Mots-clés :

Compétence, Evaluation, Faire-Cognitif, Résolution de problèmes, Tutoriel

Key-words :

Cognitive-doing, Competence, Evaluation, Intelligent Tutoring Systems, SolvingProblems

L'objectif de ce travail est l'évaluation automatique par et sur informatique de connaissances, capacités et compétences en Sciences physiques depuis le niveau le plus élémentaire jusqu'au Premier Cycle inclus des Universités. Ce travail s'intègre dans le cadre d'un programme européen LEONARDO d'Accréditation des compétences[2][8]. Il ne gère pas pour l'instant les évaluations sur supports expérimentaux ou en laboratoire .

1. Description du dispositif actuel

EvA est un tutoriel d'Evaluation Autonome assistée par ordinateur, destiné aussi bien à l'évaluation des connaissances que des compétences.
L'évaluation peut se concevoir selon différentes modalités :

a) - soit **comme un moment séparé**, se déroulant en dehors de la formation et sans lien immédiat à une source d'information. Dans ce cas, le tutoriel Eva réalise un bilan personnel de connaissances et de compétences.

b) - soit **intégrée au processus de formation**, mettant en oeuvre des liens informatifs avec la question posée. Dans ce cas, la démarche de recherche d'information et la réponse obtenue à la question posée étayent la réorganisation des représentations. Le tutoriel Eva constitue alors une aide pour l'autoévaluation et l'autoformation. Il fournit au formateur une indication individuelle et statistique sur la qualité d'une formation vis à vis d'un public donné.

La maquette d'EvA que nous avons réalisée est une unité fonctionnelle autonome.
Elle est réalisé dans un langage :

- objet, pour faciliter la conception modulaire, en différents blocs,
- hypertextuel, pour faciliter la gestion des liens entre les différents blocs.

Les différents blocs fonctionnels qui constituent EvA sont de plusieurs types :

- un bloc de gestion des évaluations : présentation, choix d'une stratégie (libre ou guidée), traitement des résultats des évaluations,
- des blocs d'évaluations (blocs de saisie) traitant de contenus spécifiques,
- des blocs de commentaires relatifs à chacun de ces blocs d'évaluation.

Dans sa réalisation actuelle, cette maquette est intégrée à un ensemble de blocs d'informations relatifs aux Sciences Physiques : le tutoriel CORREL correspondant à un contenu d'enseignement de Mécanique Elémentaire.
L'ensemble Formation (CORREL)[1] et Evaluation (Eva) est déjà utilisé régulièrement, dans le cadre d'un Enseignement de Mécanique de DEUG, à l'Université de Provence.
Les questions ou thèmes soumis à évaluation sont présentés sous différentes formes : textes, images fixes (dessin ou photo numérisée), images animées (fabriquées par simulation) ou images vidéo (obtenues par numérisation de séquences filmées).

Quel que soit le type d'information présentée (directement ou par appel d'un fichier externe), les réponses peuvent s'exprimer sous des formes différentes :

- sélection de phrases utilisées comme descripteurs d'une situation (spécifique ou générale).
- sélection d'éléments (groupes de mots ou symboles) dont l'assemblage peut constituer une phrase ou une formule complexe.
- saisie au clavier d'éléments de réponse simple, par exemple valeurs numériques.
- sélection de mots ou de groupes de mots dans un texte.
- sélection d'éléments symboliques (groupe de mots ou de symboles) dont l'assemblage peut constituer une phrase ou une formule complexe.
- désignation de points signifiants sur l'écran, par simple "clic".

Les réponses saisies par l'Usager sont analysées par le tutoriel EvA puis restituées : soit immédiatement après chaque saisie, soit en fin de session d'évaluation.

2. Le produit ForEva

D'autres blocs de traitement des évaluations sont associés à cette maquette, dans le cadre du projet LEONARDO "Modèle d'Accréditation des Compétences", en particulier:

- un bloc fonctionnel permettant d'établir une cartographie des connaissances et des compétences de son utilisateur.
- un bloc de gestion des processus pour définir des critères de pilotage des évaluations.

Les résultats de ces traitements constituent la définition du "crédit" de connaissances et de compétences.

L'objectif à atteindre est qu'une personne désireuse de s'évaluer à partir du réseau, après avoir indiqué le thème et le niveau souhaité puisse se soumettre à un guide d'évaluation spécifique (Figure 1); le traitement de ses actions et réponses donne lieu après confirmation à trois types de bilans : par connaissances, par compétences et par contexte, ce dernier étant constitué notamment des diverses modalités de présentation de la situation-problème.

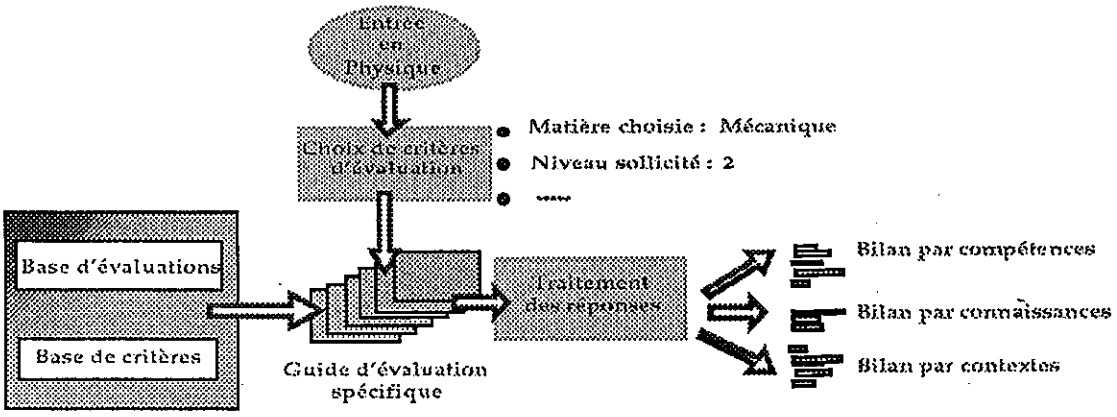


Figure 1

La hiérarchie des critères qui codent les activités d'évaluation est reproduite en Figure 2.

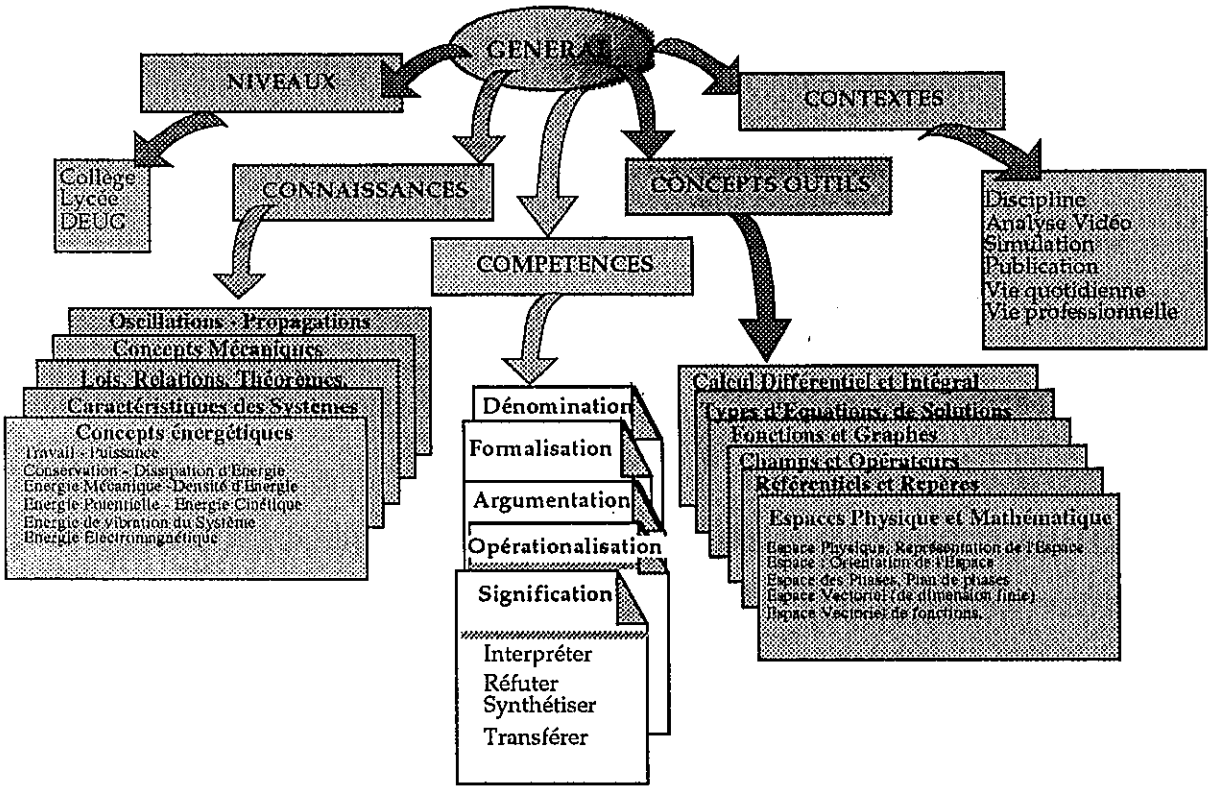


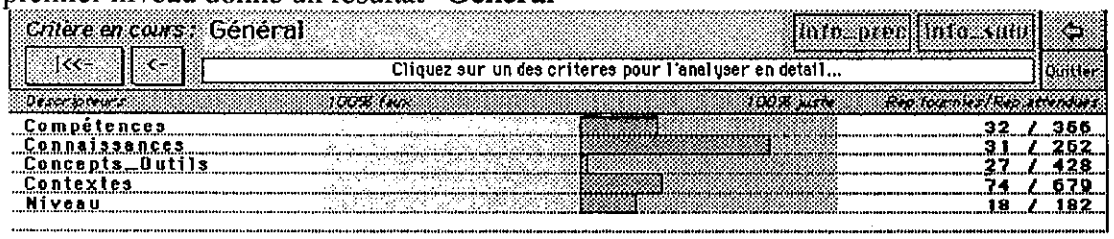
Figure 2

Les différents niveaux de bilans par critères traduisent les acquis :

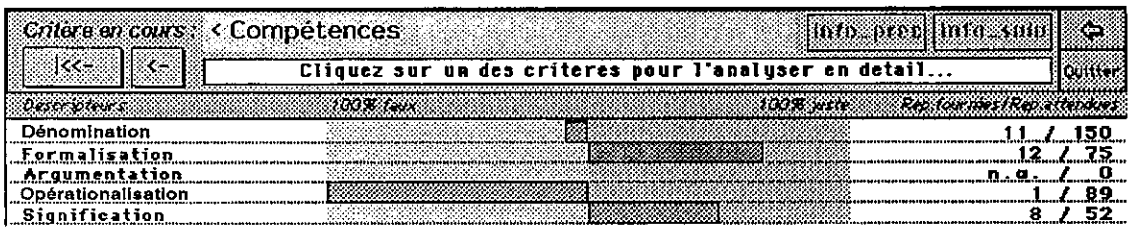
- d'un individu sur un test,
- d'un individu sur un ensemble de tests
- d'un groupe d'individus sur un test
- d'un groupe d'individus sur un ensemble de tests

Pour fixer les idées, on reproduit en Figures 3a), 3b), 3c), 3d) les bilans-résultats d'un groupe d'étudiants de DEUG obtenus en situation présentielle, dans un dispositif d'auto-formation, à propos d'une activité d'évaluation sur ordinateur mettant essentiellement en jeu les compétences regroupées sous les rubriques : Formalisation, Résolution, Signification à propos de concepts de cinématique et de mécanique des vibrations.

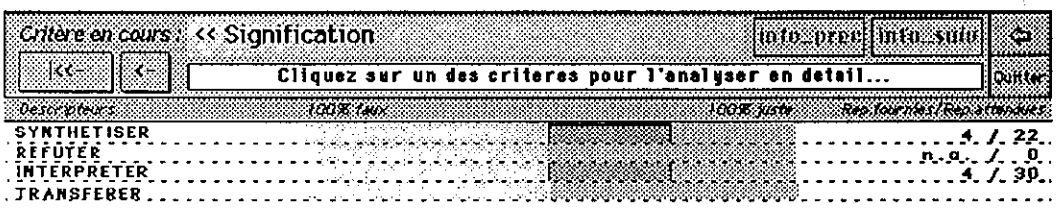
Le premier niveau donne un résultat "**Général**"



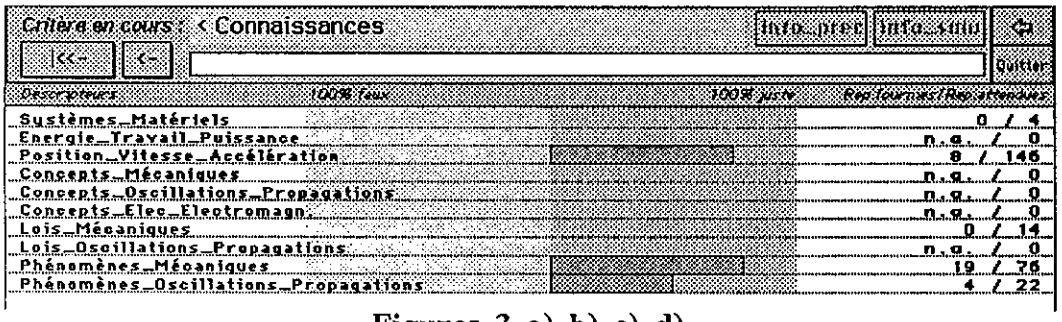
En sélectionnant la rubrique "**Compétences**" , on obtient le bilan des résultats des grands groupes de compétences identifiées.



En sélectionnant la rubrique **"Signification"**, on obtient le troisième niveau de bilan, plus précis, concernant les activités mentales requises dans cette évaluation : **Synthèse, Réfutation, Interprétation et Transfert.**



Si l'on revient à l'écran "Général", la sélection de la rubrique "Connaissances nous indique la répartition des résultats selon les items de la session d'évaluation. Un troisième niveau est également accessible sur les items de connaissance.



Figures 3 a) b) c) d)

Les divers niveaux des différents critères sont reportés sur la Figure 2

La traduction d'EVA, pour la mise en oeuvre d'une évaluation sur réseau (réseau interne et Internet) est en cours de réalisation.Le nouveau produit ForEva reprend en grande partie les fonctionnalités du logiciel CORREL-EVA. Les principales nouveautés de ce développement se situent au niveau de la mise à disposition sur Internet des informations et de la possibilité d'exécuter les processus d'évaluations à distance. De plus nous prévoyons d'intégrer et d'améliorer la gestion de stratégies d'évaluation et de remédiation.

3. Les bases théoriques

Le projet ForEva comporte plusieurs aspects:

- informatique avec la réalisation du produit,
- pédagogique avec la conception de situations-problèmes donnant lieu à évaluation,
- documentaire et au-delà sémiotique avec l'identification des connaissances et compétences puis leur codage,

- épistémologique enfin car l'évaluation de compétences en sciences (ou dans dans tout autre domaine donnant lieu à théorie et formalisation) ne saurait se passer de points de vue sur la structure et le fonctionnement du domaine concerné.

Bien évidemment ces quatre aspects sont articulés dans la réalisation du produit : le cahier des charges informatiques traduisant les contenus, les options pédagogiques et épistémologiques à l'aide notamment des repérages et catalogage sémiotiques.

Nous ne parlerons ici que des deux derniers aspects car il semble que leur prise en compte effective et réalisée (et non virtuelle!) dans un système informatique constitue une innovation.

Traditionnellement, l'évaluation en sciences physiques comporte deux modalités:

- l'une, énonciative, de l'ordre de la simple **transmission d'un savoir acquis** permet d'évaluer les connaissances manifestées dont on suppose qu'elles sont stockées chez le sujet, si toutefois les conditions de passage du test autorisent cette inférence. Il s'agit alors d'un acte répétitif qui permet seulement de supposer au préalable la compétence d'énonciation,
- l'autre modalité la plus couramment utilisée est la **Résolution de problèmes.**

Remarques

1°) L'utilisation de l'expression **"Résolution de problèmes"** doit s'entendre au sens cognitif et regroupe en fait toute la catégorie d'activités qui part d'une situation porteuse d'interrogations et qui pose de ce fait problème. A ce titre, cela concerne une gamme bien plus large d'activités que celles intervenant en Physique, et même au-delà, en sciences formelles. Cela peut concerner des activités de maintenance, de recherche de pannes, d'adéquation d'un produit à des besoins implicites, et se rattacher ainsi à des pratiques professionnelles(6) (7).

2°) Les conditions de l'évaluation liées aux contrats -social et pédagogique- avec les Institutions (Destinateur/Destinataire) (4) ni les modalités d'adhésion/rejet aux procédures de sanction/récompense, fondamentales dans les processus d'acquisition ne sont pas pris en compte dans cette phase du travail : le sujet contractualisé est supposé désireux de s'évaluer...

• S'il s'agit de problèmes résolus et bien assimilés auparavant, en phase d'apprentissage par le sujet, les savoir-faire nécessaires existent déjà et sont reproduits : la séquence d'évaluation est alors la répétition puis la manifestation par le sujet des actes mentaux acquis lors de la formation.

• S'il s'agit d'un problème nouveau, la modalité d'évaluation engendrera un véritable processus de création de la part du sujet.

Mais acquis dans l'apprentissage ou créés pour l'occasion, ces savoir-faire reproduisent toujours tout ou partie de la méthode scientifique générale et de l'activité de modélisation inhérente à la plupart des situations formalisables, ancrées sur la "réalité" des situations.

Il semble donc nécessaire de se doter d'un modèle maximal de l'activité de résolution de problèmes sachant qu'une séquence concrête d'évaluation n'en comportera souvent qu'une partie.

4. Le modèle de l'activité de Résolution de problèmes

On suppose que le savoir scientifique, préexiste à l'activité du sujet en évaluation et comporte un certain nombre d'entités telles que concepts, méthodes, théorèmes, lois....mais aussi procédures de résolution d'équations, de calcul, de mesure...
Nous dirons qu'un énoncé est en langage naturel par opposition au langage scientifique, - qu'il soit conceptuel ou mathématique -, lorsqu'il ne comporte que des termes relevant de la forme la plus simple d'un premier niveau de langage : le naturel, à l'image du concret, est toujours relatif à l'état de l'étape antérieure.
De même, nous parlerons de situation "naturelle" pour désigner un énoncé en langage naturel sur la situation ou un film sans commentaires montrant un phénomène.

Nous appellerons réduction ou **focalisation** l'opération qui permet à partir d'une situation complexe, multidimensionnelle, de focaliser sur une dimension redevable du champ concerné par le projet d'évaluation : on évalue toujours quelque chose pour une instance. L'élimination de dimensions non pertinentes n'aboutit pas forcément au bon résultat lors du premier examen des données : l'âge du capitaine peut bien intervenir sur la longueur du bateau si la courbe donnant la relation entre les années d'expérience et la taille des navires à commander fait partie de la troisième question....
Cette focalisation rend compte de la manière dont s'effectue la sélection de certaines données aux dépens d'autres dans le projet de fabrication d'objets théoriques que constitue, à partir d'une vision "naturelle" décrite avec des mots ou montrées par une vidéo, la détermination de l'ensemble pertinent de données : mots, propriétés, représentations du fonctionnement.

L'opération de subsomption(4) qui permet d'aboutir depuis les données initiales au problème scientifique qui le modélise se décompose en plusieurs :

- l'**identification** qui permet d'obtenir à partir des entités et de leurs propriétés les concepts physiques correspondants,
- la **projection** comme acte cognitif par lequel le sujet agence les concepts.

Lorsque projection et identification permettent une **interprétation** de l'énoncé qui infirme la cohérence de l'ensemble des données, il faut alors "réviser sa position", effectuer une opération d'**ajustement** qui modifie la première représentation; cette opération présuppose une compétence d'appréciation du degré de cohérence de l'ensemble des données.

La **mise en adéquation** du savoir théorique au problème identifié peut entrainer le besoin de **différenciation** entre deux concepts, entre deux méthodes; l'**énonciation** des éléments théoriques retenus pour l'**application** est alors indispensable.

La résolution de problèmes n'étant en fait qu'une série de transformations de l'énoncé en sa solution, les transformations d'entités prendront une place importante : **résolution** au sens strict des mathématiques (équations), **calcul** ou mise en évidence d'une propriété à l'aide d'une **transfiguration** des données; ces transformations sont la plupart du temps dictées par l'**argumentation** (- qui peut être de nature propositionnelle ou hypothético-déductive).

La réponse au problème ne peut se limiter à la découverte de la solution dans l'espace de description : elle ne devient vraiment solution puis réponse qu'après un double processus d'**interprétation** et de **réfutation.** L'établissement de la signification peut également comporter une phase de **transfert** vers d'autres situations : c'est quelquefois seulement à l'issue de cette dernière opération que s'élabore la signification de la solution, donc celle du problème.

L'ensemble des fonctions intervenant dans notre modèle de résolution s'organise donc en cinq phases reprises ci-dessous:

1°) Dénomination de l'espace pertinent de données	
IDENTIFIER-REPERER	les concepts mis en jeu avec leurs propriétés théoriques.
REDUIRE-EXTRAIRE	la "complexité" d'une situation réelle en sélectionnant concepts, propriétés et représentation du fonctionnement.
ENONCER	une propriété, une connaissance.
2°) Formalisation	
AJUSTER	la représentation à l'ensemble des données du problème.
DIFFERENCIER-RELIER	concepts, arguments.
APPLIQUER	le savoir théorique au problème posé.
INDUIRE-PREDIRE	une propriété, un résultat.
3°) Déploiement de l'argumentaire	
CHOISIR	une méthode avec justification logique des choix.
ARTICULER	des arguments. Déduire, établir, démontrer.
4°) Opérationalisation	
RESOUDRE	Transformer en éléments plus simples par des lois de composition internes.
CALCULER	Passer de données ou de résultats littéraux aux mêmes mais numériques.
MESURER	Obtenir une valeur numérique pour un concept.
TRANSFORMER	un type de donnée ou de résultat en un autre sans en modifier la nature profonde.
5°) Signification	
SYNTHETISER	un ensemble d'expériences, de résultats...
INTERPRETER	grandeur physique ou résultat en termes "naturels".
REFUTER	en démontrant la fausseté d'une réponse trouvée.
TRANSFERER	d'un espace de situation à un autre, à un autre modèle...

5. Le statut des opérations

Si nous comparons la résolution d'une énigme à la création d'une oeuvre, il nous faut repérer non seulement l'évolution des ébauches successives au cours du processus de création mais également les transformations que subit l'auteur lorsqu'il fabrique mais aussi lorsqu'il met en perspective, rectifie ou juge d'un effet.
Seule, l'oeuvre est manifeste; quelquefois, l'auteur décrit, ailleurs, en marge, ses hésitations, ses modifications de stratégie.

De même que la statue est déjà présente dans le bloc de pierre, la solution d'une énigme est présente dans son énoncé initial, faute de quoi la compétence de l'énonciateur serait contestable.
L'objet d'études subit donc une série de transformations, avec l'aide de savoirs théoriques déjà là et de langages appropriés.
Le sujet, non seulement réalise ces transformations, mais également s'en approprie le résultat, effectuant ainsi des performances qui doivent être manifestes pour l'évaluation.
Mais le sujet de faire ne peut réaliser la transformation que s'il possède au préalable la compétence, condition nécessaire à la réalisation de l'acte; la compétence est le *Ce qui fait être performant* dans un contexte donné.

Une véritable prise en compte des compétences d'un sujet ainsi définies risque fort de passer sous la critique d'un mentalisme trop poussé!
Par contre, les compétences du produit d'évaluation définies comme "l'ensemble des opérations qui permettent au sujet de faire ces transformations" peuvent s'étudier : c'est le cahier des charges informatiques qui se doit de médiatiser les compétences cette fois pédagogiques, exercées en amont par le concepteur-énonciateur des contenus d'évaluation.

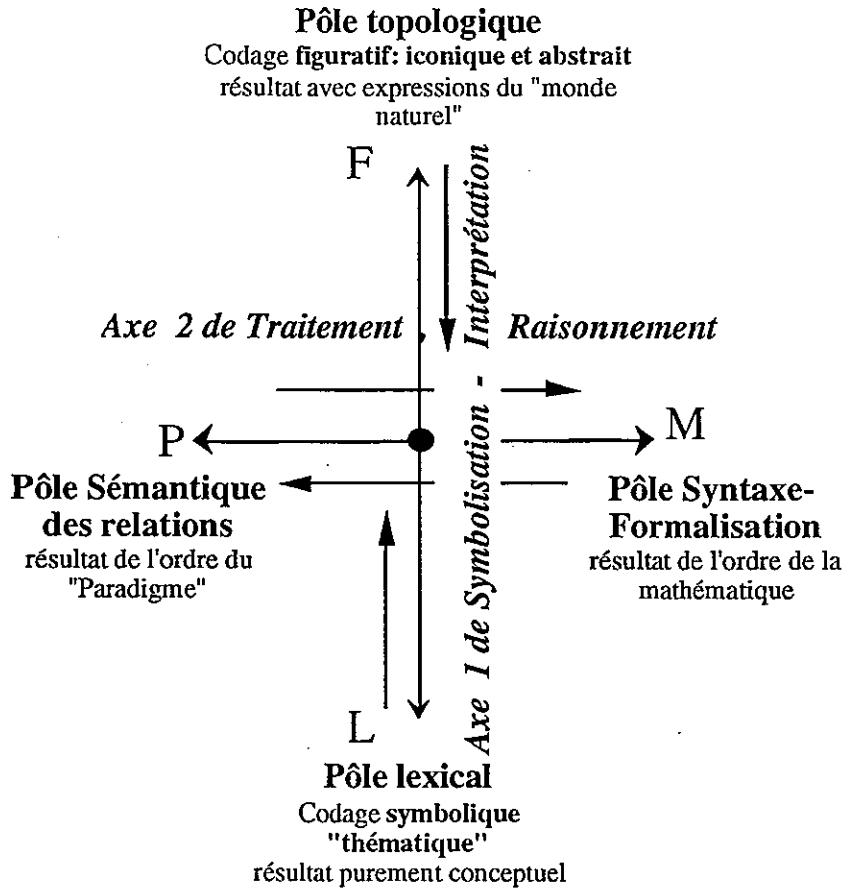
Nous serons amenés à considérer qu'au cours de l'évaluation de la résolution de problèmes :

- les différentes opérations sont manifestables (ce qui relève de la compétence du concepteur) et manifestées sous forme de signes-réponses (inscrits sur des écrans, des copies...) par le sujet à évaluer ,
- le sujet producteur de l'opération, a auparavant, bien acquis la compétence qui lui permet de la réaliser (ce qui relève de la mise en place des conditions concrètes du test d'évaluation : pas d'informations piratées !...).

Alors, (et l'hypothèse est de taille!), l'examen des performances manifestées permet d'inférer la présence de compétences actualisées dans la performance.
Encore faut-il que le concepteur du produit d'évaluation ne se cantonne pas aux manifestations usuelles seulement répétitives et produise des énoncés d'énigmes qui rendent visibles l'ensemble des transformations retenues pour participer à la résolution.

Si nous reprenons la métaphore de la statue, le faire transformateur de l'objet est tangible: la $n^{ième}$ ébauche est différente de celle d'avant ou d'après.
Dans notre cas, les faire-transformateur portent sur des objets de savoir, investis de valeur cognitive liée au savoir sur chaque objet; il s'agit donc au sens de la sémiotique de faire-cognitif(3).
Par abus de langage, nous appellerons **fonctions ces "faire-transformateur" {F} portant sur les objets cognitifs que sont les connaissances.** Elles tiennent lieu de compétences dans notre système d'évaluation. Une fonction F appliquée à un argument x, permet la performance F(x) qui, manifestée, pourra donner lieu à évaluation.

L'ensemble des fonctions se distribue sur deux axes (5) reliant quatre pôles :



- l'axe 1 de **Symbolisation-Interprétation** reliant le pôle F figuratif, - iconique ou abstrait- au pôle lexical L des concepts,

- l'axe 2 de **Traitement-Raisonnement** reliant le pôle sémantique P des relations physiques au pôle M mathématique.

6. Perspectives

L'étude détaillée des propriétés de ces fonctions est actuellement en cours; concrètement, cette étude débouche nécessairement sur les **procédures de codage-décodage** effectuées par le concepteur-évaluateur. Elle détermine en conséquence, un ensemble de contraintes portant sur les contenus d'évaluation qui doivent permettre la manifestation des différentes fonctions sélectionnées par notre modèle.

7. Bibliographie

(1) COSTE J.P. "Hypertextes et stratégies pédagogiques" Actes du Colloque: Les nouveaux outils pour l'Enseignement Supérieur" Marseille Juin 1993

(2) COSTE J.P. "Modèle d'accréditation des compétences" Contrat LEONARDO LdV F/96/21498

(3) COURTES J. "Analyse sémiotique du discours" Hachette 1991

(4) GREIMAS A.J. COURTES J."Sémiotique" Tomes I et II Hachette 1986

(5) MANDELSOHN P. "Les capacités transversales" Conférence Université d'Eté " Les concepts-clés du formateur" MARSEILLE juillet 1993

(6) MANDON N."Analyse des emplois et gestion anticipée des compétences" Education Permanente n° 109 - Déc 1990 p.101-109

(7) MICHEL S. LEDRU M. "Description des compétences et formation, une approche cognitive" Education Permanente n° 109 - Déc 1990 p.109-115

(8) REIFFERS J.L. Rapport d'Etape du Projet LEONARDO "Accréditation des Compétences" mars 1998

VOUZZAVEDIBISAR,

La chanson du Mammouth.

C'était un gros mammouth - mouth - mouth,
qui marchait dans un pré - pré - pré,
il y avait plein de fleurs - fleurs - fleurs,
il a tout piétiné - nez - nez.

Au ciné de l'école,
c'était le gros mammouth
qui vendait le programme.
Le film était mauvais,
il racontait l'histoire
d'un cancre paresseux,
qui finit au chômage.
Et puis de son copain
le brillant lauréat
qui trouve rien non plus
et finit chemineau.

C'était Miraculé, laid, laid,
avec casque et matraque - traque - traque,
traque les sans papiers - pied - pied,
qui traque et se détraque - traque - traque.

Au ciné de la rue,
c'était Miraculé
qui vendait le programme.
Le film n'était pas bon,
ça se passait là-bas,
au fond, dans un ghetto,
des jeunes s'ennuyaient,
et pour passer le temps
ils brûlaient des autos,
cassaient quelques vitrines,
finissaient en prison.

Le chef des percepteurs - teur - teur,
fabriquait des impôts - pot - pot.
A tort et à travers - vers - vers,
jetait tout dans le gouffre - bouffre - bouffre.

Au ciné de la bourse,
c'était le percepteur
qui vendait les bonbons,
le film était minable
c'est les marchands de sous
qui remportaient les prix
et planquaient leur galette
dans les verts paradis
dont l'argent refleurit
dans les jardins rêvés,
où pousse l'arbre à came.

Le gros mammouth - mouth - mouth,
qui marchait dans les prés - prés - prés,
c'était tout plein de fleurs - fleur - fleur,
fallait le dégraisser - tsé - tsé.

Au ciné du travail,
c'était le gras mammouth
qui vendait les billets,
comme au ballet pourri
des chaises musicales
on payait les violons
mais il manquait des places,
l'argent était parti
acheter des machines,
et puis surtout des armes,
et aussi des ministres,
et puis des Présidents.

C'était un gros mammouth - mouth - mouth,
qui marchait sur l'école - colle - colle,
il voulait que l'élève - lève - lève,
soit côté à la bourse, marche et crève.

Edmond Bianco

Université de Provence
Atelier de Reprographie
Centre Saint Charles
3, place Victor Hugo
F - 13331 Marseille Cedex 3