

Le système HALPIN : recherche documentaire en langue naturelle et dialogue multimodal

José Rouillard¹ et Jean Caelen

Laboratoire CLIPS IMAG BP 53 38041 Grenoble Cedex 9 - *E-mail* : Jose.Rouillard ; Jean.Caelen {@imag.fr}

Résumé : Le système HALPIN permet à un utilisateur du World Wide Web de dialoguer en langue naturelle avec une machine distante afin d'effectuer une recherche documentaire (sur la base de données documentaire de l'INRIA). Le système a été conçu dans une démarche incrémentale à partir d'une enquête d'usage effectuée auprès d'utilisateurs d'une médiathèque qui a permis ensuite de formaliser un modèle de dialogue homme-machine basé sur la reconnaissance de concepts dans les énoncés. Le modèle lui-même s'appuie sur un graphe de tâche qui coordonne les différents sous-dialogues. Le logiciel développé met en œuvre des composantes de reconnaissance automatique de la parole, de dialogue personne-système, de recherche d'information, et de synthèse vocale. Ce logiciel a été évalué avec un panel d'utilisateurs : les résultats montrent une bonne acceptation du dialogue vocal pour la recherche d'information ainsi qu'une bonne complémentarité avec une stratégie de navigation.

1. Position du problème

La recherche d'information est une question qui prend de l'importance avec le Web, du fait de la dissémination et de la fragmentation des informations dans de très nombreux sites. D'un site à un autre, ces informations ont souvent peu de rapport entre elles (malgré les liens hypertextuels qui semblent les associer fortement). Il est donc pratiquement impossible d'effectuer une navigation thématique puisque les documents ne sont pas indexés, sauf en utilisant un moteur de recherche généraliste pour "ramener" à l'avance un certain nombre de pages lexicalement liées, et de les parcourir ensuite dans son propre espace de travail. Or sur un plan cognitif, l'utilisateur ne procède pas du tout ainsi : il a plutôt tendance (a) à *butiner sur un fil thématique* s'il n'a pas de contrainte de recherche précise ou de but particulier à satisfaire autre que sa propre curiosité ou (b) à vouloir *accéder rapidement aux informations* dont il a besoin pour un travail donné. Dans le premier cas, l'aide que peut lui fournir un système interactif sera de *l'accompagner* dans son parcours thématique (informations sur le *où* et le *quand* obtenir : aide à la navigation, au repérage, à la collecte, etc.), dans le deuxième cas, l'aide sera plutôt de lui fournir les conditions optimales *d'accès* (informations sur le *quoi* et le *comment* obtenir).

Le système HALPIN que nous développons au sein des projets Orion [ORI 99] et NTI-SPI-Santé vise le deuxième cas d'utilisation, pour la recherche de documents *structurés* dans un premier temps, puis plus largement sur le Web plus tard. Les hypothèses de travail que nous posons sont :

- Le système se comporte comme un assistant dans la tâche de recherche documentaire, il ne doit pas être trop intrusif mais suffisamment coopératif,
- L'utilisateur peut dialoguer avec lui de manière multimodale (parole, écrit, geste de désignation), le niveau de langue ne devant pas être trop contraint,
- Le système a la capacité de s'adapter à son interlocuteur (niveau d'habitude, handicaps éventuels).

Nous avons choisi de nous limiter aux documents structurés, pour éviter de faire une analyse de contenu des documents parcourus. Cette limitation n'est pas un obstacle à la généralisation du système pour tous types de documents ou de pages Web. L'utilisateur se trouve, dans Halpin, face à l'écran de l'ordinateur (Figure 1), ce dernier connecté au réseau Internet. Il dispose d'un clavier/souris et de dispositifs audio pour la parole (microphone, enceintes acoustiques, carte son et éventuellement amplificateurs d'entrée et/ou de sortie).

¹ Coordonnées actuelles de cet auteur : Laboratoire Trigone - CUEEP Bât B6, Université des Sciences et Technologies de Lille 59655 Villeneuve d'Ascq Cedex – France - *E-mail* : Jose.Rouillard@univ-lille1.fr



Figure 1. L'utilisateur dialogue avec le système HALPIN. Il énonce (dans cet exemple) sa demande oralement et le système lui répond sur le même mode. La réponse est coopérative : c'est ici une proposition d'affiner la requête pour restreindre le nombre de documents trouvés.

2. Relations à l'éducation et à l'enseignement

Les changements apportés par l'existence du Web affectent aussi le monde de l'éducation : 10000 établissements scolaires sont connectés en France – même si la plupart n'ont qu'un seul poste relié à l'Internet. Le Web offre de formidables ressources pour l'enseignement, tant par ses contenus que par sa possibilité de délivrer des cours à distance. Les enseignants qui ont tenté l'aventure en arrivent tous à la même conclusion : c'est intéressant, mais... « Il reste à construire des outils intellectuels préalables basés sur la représentation mentale de l'information et sa structure (recherche par association d'idées, par analogie, par induction, etc.) qui favoriseraient la manipulation de l'information et son intégration et dans un deuxième temps la forme affective de la communication, l'échange. » [BUC 98]. Cependant, les élèves préparent leurs dossiers sur l'ordinateur domestique (1,25 millions fin 98). L'enseignement ouvert et à distance se développe, les pourvoyeurs de formation se diversifient – ce sont parfois les entreprises elles mêmes (ex. l'Université Microsoft). Les nouveaux outils de la communication médiatisée par ordinateur favorisent le travail collaboratif à distance et la relation entre les différents acteurs de la formation.

Dans ce contexte, apprendre à maîtriser les techniques d'accès à l'information, est devenu une priorité. Les usagers développent leurs compétences mais les systèmes se complexifient et la « surcharge d'information » menace : les aides à la recherche et à la collecte d'informations paraissent aujourd'hui plus nécessaires qu'hier. Une des préoccupations d'actualité est celle d'une « société de l'information » où les systèmes techniques risqueraient d'accentuer – plutôt que de réduire - les inégalités dans l'accès au savoir. L'enjeu pour ces systèmes est donc clairement d'aider les usagers à maîtriser et réussir leurs tâches de recherche d'information.

C'est un défi qui ne se pose pas qu'en termes purement techniques : quelles sont les capacités cognitives des usagers, quels sont leurs buts, leurs stratégies, comment des logiciels peuvent-ils les aider à progresser vers ces buts et les aider à décider qu'ils les ont atteints, quelles sont les informations « qui comptent » [HAR 98] pour eux et pour les institutions au sein desquelles ils travaillent ? Voilà les questions qui guident le processus de conception des dispositifs d'Halpin pour la recherche d'information à l'adresse des étudiants. Si les solutions proposées exploitent les avancées techniques les plus récentes comme l'analyse du langage naturel, l'utilisation de composants logiciels ou le recours à des services de traitement en ligne, elles restent fondamentalement au service des usagers – en tant qu'individus pris dans un contexte institutionnel d'usage. En bref le processus de conception n'est pas techno-centré, mais centré sur l'usage [BEL 97] et l'analyse de l'usage linguistique est ici un aspect important du projet.

3. Navigation versus dialogue

Déjà avec les corpus hypermédia « fermés » des années 80, les recherches avaient mis en évidence des difficultés spécifiques de la navigation [CON 87]. Ces problèmes ont pris une nouvelle dimension avec l'arrivée – au début des années 90 – du Web. On peut cependant les ramener à deux grands types de problèmes : la désorientation et la surcharge cognitive. Chaque type de problème a été l'objet de nombreuses études spécifiques [CAS 96], [SOU 96], [CAR 96].

La **désorientation** provient de l'absence de repères des usagers lorsqu'ils parcourent les réseaux : ils ont besoin de savoir d'où ils viennent, où ils sont et comment se rendre d'un endroit à l'autre. Trois niveaux de difficultés ont été mis en évidence : 1) la navigation dans l'espace logiciel - c'est le fonctionnement du système hypertexte qui n'est pas connu, l'utilisateur ne sait pas ce qu'il faut faire, 2) la navigation dans l'espace conceptuel - ce sont les catégories qui organisent l'information qui ne sont pas familières, l'utilisateur n'arrive pas à rentrer dans le système conceptuel de l'auteur, 3) la navigation dans le texte - c'est le cheminement dans les liens hypertextuels qui n'est pas familier, l'utilisateur perd le fil de sa navigation. (Cf. [BAL 96] pour plus de précisions).

La **surcharge cognitive** est liée aux nombreuses décisions que doit faire l'utilisateur lorsqu'il parcourt un hypermédia : quels liens suivre, comment retrouver ceux qui l'intéressent dans ceux qu'il a parcourus ou dans ceux qu'il lui reste à parcourir, etc. La lecture hypertextuelle s'avère hautement spécialisée et requiert des schèmes de lecture spécifiques. « L'utilisation d'un système d'hypertexte suppose que le lecteur soit en mesure de gérer son propre parcours dans l'information » ([ROU 97] p. 165). On constate un manque de stratégies opératoires, une perte des repères discursifs normalement présents avec le texte imprimé. Existe-t-il des modes de lecture/compréhension, et des stratégies, qui seraient spécifiques aux hypertextes ? L'utilisateur doit pouvoir trouver l'information qu'il recherche en naviguant à travers des liens et des nœuds. Ses tâches de recherche d'information exigent qu'il accède 'intelligemment' à l'information pertinente: ce qui suppose des capacités telle que aller d'un endroit à un autre, identifier le document atteint, l'évaluer, le ranger ou mémoriser son adresse, et faire les liens avec d'autres documents et informations. Castelli et al. [CAS 96], ont identifié six variables cognitives clés dans la navigation en hyper-espace : efficacité intellectuelle, capacité analytique, flexibilité analytique, capacité de synthèse, raisonnement abstrait, indépendance du champ cognitif. La navigation s'avère un processus complexe.

Pour résoudre ces problèmes de navigation, on a d'abord cherché des solutions du côté d'une re-structuration des corpus. Si dans les systèmes hypermédia fermés, on a pu améliorer la conception du réseau de nœuds et liens dans lesquels les usagers naviguent, ces solutions ne s'appliquent que très partiellement au Web : on peut en effet structurer un site institutionnel par exemple (parce qu'il est géré par un modérateur local) pour y faciliter la navigation, mais on ne peut pas imposer de structure au réseau des liens qui conduisent à ce site. Les logiciels de navigation tel Navigator de Netscape ou Explorer de Microsoft proposent des fonctionnalités comme l'historique et les « bookmarks » ; mais ces procédés techniques s'avèrent être des aides insuffisantes pour pallier aux difficultés de l'utilisateur. De plus les représentations que se font les usagers d'un système hypertextuel dans son ensemble, ou des hypermédiats spécifiques qu'ils parcourent, sont très variables d'une personne à une autre. Dans ce contexte les solutions passent sans doute par un développement des outils techniques et des compétences des usagers. C'est cette perspective qu'adopte le projet Halpin en proposant une « instrumentation » [RAB 95] de l'activité dynamique et évolutive de recherche d'information qu'implique l'usage de l'hypermédia et du Web. La navigation ne se fera donc plus seulement de manière traditionnelle, en manipulant le clavier et la souris pour piloter un moteur de recherche, mais également par le biais d'un dialogue en langue naturelle avec la machine, tout au long de l'activité de recherche documentaire.

Qu'apporte le dialogue ?

L'approche dialogique se distingue de la navigation dans la mesure où l'on formule directement ce que l'on désire : l'accès est plus immédiat si l'on sait formuler sa demande (par rapport à la navigation qui consisterait à faire du porte à porte pour rechercher une personne et à sonner pour voir qui répond, la métaphore serait ici de crier dans la rue pour savoir si cette personne y habite). Le problème n'est plus de se repérer mais de respecter et de maîtriser le code de communication (mots, énoncés, conventions linguistiques, etc.). On peut supposer qu'après un rapide apprentissage, ce sera le cas pour un dialogue en (LN) langue naturelle (ou un sous-langage spécialisé). On attend donc de ce dernier, une diminution de la charge cognitive par une prise en charge de l'interaction et de la coopération [ROU 98a], afin de :

- rendre l'interaction plus souple par l'usage du langage naturel et des représentations graphiques,
- permettre un guidage efficace des utilisateurs par focalisation sur un thème ou un sous-thème donné, par élargissement ou affinement, thématique,
- permettre la formulation de requêtes imprécises : un nom d'auteur incomplet, une date approximative, etc.
- permettre l'usage de procédés tels que la reformulation, l'auto-correction, la clarification, etc. qui allègent la charge cognitive de l'utilisateur dans la mesure où il peut continuer à réfléchir tout en formulant sa demande,
- permettre un retour simple et direct aux documents déjà visités qui allège la charge cognitive de l'utilisateur en remplaçant un processus de rappel mémoriel par un processus de « dialogue ».

En Dialogue Homme-Machine (DHM), la robustesse des modèles de dialogue ainsi que la confrontation des prototypes à une utilisation réelle (hors laboratoire) demeurent des éléments importants pour valider toute nouvelle approche du problème. En effet, un système de DHM en LN doit pouvoir analyser et générer des énoncés aussi bien à propos de la tâche elle-même que du dialogue à propos de la tâche. En d'autres termes, un système sera jugé non seulement sur ses capacités dialogiques intrinsèques, mais également sur ses aptitudes à donner du sens à ce que dit son interlocuteur. A la question « Avez-vous l'heure ? », un système performant doit comprendre qu'il ne suffit pas de répondre par oui ou par non. Cela implique une certaine connaissance du monde et des objets qui le composent, pour une tâche donnée.

4. Le système Halpin

Le système Halpin (Hyperdialogue avec un Agent en Langage Proche de l'Interaction Naturelle) utilise les deux modalités de recherche d'information : par navigation, et par dialogue (écrit ou oral). Il joue d'une part le rôle d'un bibliothécaire qui sait *où* se trouve un ouvrage, sait *comment* aller le chercher, mais ne sait pas ce qu'il contient et il joue d'autre part le rôle d'une bibliothèque en libre service où il est permis de butiner soi-même. Il n'a pas de compétence sur le domaine ou le thème de l'ouvrage et ne vise pas à en acquérir, bien que de telles connaissances soient certainement utiles dans sa tâche de recherche. Le système Halpin n'est pas non plus un documentaliste qui saurait réunir, de lui-même, une liste d'ouvrages en vue d'une étude déterminée ou sur un thème précis. Il n'a pas la mémoire de recherches passées.

4.1. Description

Le système Halpin est essentiellement un *assistant dans la recherche immédiate et dans la collecte* : il permet à un utilisateur du Web de dialoguer en langue naturelle avec une machine afin d'effectuer une recherche documentaire (pour le moment dans la base de données documentaire de l'INRIA). L'interaction est multimodale : l'utilisateur tape ou prononce son énoncé, en français, dans une boîte de dialogue. La machine l'analyse et répond de manière vocale et textuelle. Les stratégies de dialogue s'articulent en fonction des compétences supposées (novice ou expert) et du but de l'utilisateur (retrouver un document connu ou rechercher un ensemble de documents sur un critère donné). Le système intègre des modules de reconnaissance automatique de la parole, de dialogue homme-machine, de synthèse de la parole et un moteur de recherche d'information.

L'interface du système Halpin est divisée essentiellement en deux parties : à gauche les zones de dialogue, à droite les zones de navigation. L'interface graphique du système s'utilise dans une fenêtre d'un browser classique (comme Internet Explorer de Microsoft par exemple). Elle est divisée en fenêtres ou zones : l'historique du dialogue permet de suivre l'évolution du dialogue sous forme écrite (on peut réécouter le dernier message vocal), la boîte de saisie de l'énoncé courant (que l'on peut remplir par écrit ou laisser à la charge du module de reconnaissance vocale), les boutons de réglage de l'interface vocale, les deux zones de présentation des résultats (la liste des documents d'une part et le détail en-dessous d'autre part). Ainsi cette interface est doublement multimodale, (a) au sens classique du terme, elle offre des entrées-sorties écrites, graphiques et parlées, et (b) au sens procédural elle permet de dialoguer ou de naviguer.

4.2. Limites et contraintes d'utilisation

Les contraintes d'utilisation sont liées aux médias d'interaction : parole et/ou clavier. Il est certain que la qualité du module de reconnaissance vocale est primordiale ; nous avons choisi le système ViaVoice d'IBM pour des raisons de disponibilité sur le marché. Ce système est plutôt conçu pour la dictée vocale, il ne remplit donc pas toutes les exigences nécessaires au dialogue homme-machine (il interdit par exemple l'élocution spontanée). Il est cependant d'un niveau de performance acceptable après une période d'apprentissage du locuteur. On peut également enrichir son vocabulaire pour une tâche particulière.

Une autre limite réside dans le temps nécessaire au système pour formuler une réponse aux énoncés des utilisateurs. Nous verrons plus loin que compte tenu des différents traitements effectués et des débits actuellement disponibles sur les réseaux, ce temps d'attente n'est pas si élevé qu'il y paraît. En l'espace de quelques secondes, le système analyse syntaxico-morphologiquement l'énoncé (sur un serveur distant), fait émerger les concepts reconnus, interroge la base documentaire de l'INRIA (toujours à distance), analyse les résultats obtenus, prépare une réplique adaptée (selon le but, le type d'utilisateur, la stratégie en cours, l'historique, etc.), synthétise vocalement cette réponse, et l'envoie à son destinataire en même temps que les messages (réponses textuelles et hyperliens) destinés à l'affichage graphique pour le navigateur. Nous sommes évidemment dans un système multi-utilisateurs, ce qui signifie que le système Halpin gère simultanément plusieurs connexions avec des interlocuteurs différents. Ce n'est pas tant le transit des messages vocaux qui ralentit le flux des échanges entre les allocutaires, mais bien le calcul de la synthèse vocale « à la volée ». En effet, les messages synthétisés ne sont pas préenregistrés, mais calculés en temps réel, et intègrent des éléments fournis par les usagers, pour répondre de manière naturelle et conviviale (exemple : J'ai trouvé 5 documents ayant comme auteur Turing. Consultez-les sur la page de droite. Voulez-vous affiner votre requête ?). De fait, les utilisateurs ne prennent pas forcément conscience de ces aspects techniques, et jugent souvent l'interaction avec la machine comme étant trop longue (cf. partie 6.2.) et par transitivité, peu naturelle, puisque moins spontanée.

Enfin, une autre limite technique a été constatée : la puissance d'expression du langage naturel surpasse souvent les capacités du système opératoire, de sorte que certaines commandes sont parfaitement interprétées par le système Halpin mais impossible à exécuter au niveau de la base de données. C'est le cas avec ces quelques exemples extraits du corpus Halpin :

- « Imprime moi les 3 dernières listes de résultats » : (cause : impossible de déclencher une impression sur une imprimante distante).

- « et bien alors donne moi tout depuis 75 » : (cause : l'interrogation par requête à la base de données de l'INRIA ne prévoit pas la formulation de fourchettes de dates comme 1975-1999).

- « Tu aurais pu faire un tri au moins » : (cause : impossible d'effectuer un tri de pertinence des résultats obtenus dans la base de données de l'INRIA).

Dans ces circonstances, le système explique à l'usager que sa phrase a bien été comprise, mais que l'opération correspondante ne peut pas être effectuée.

4.3. Méthode de conception

La méthode de conception d'Halpin suit un cycle de développement en spirale : après analyse de l'usage et des fonctionnalités on développe une maquette que l'on améliore au fur et à mesure des tests. Cette démarche est d'autant plus pertinente ici que pour réaliser des systèmes interactifs, les chercheurs du domaine manquent de corpus de dialogues homme-machine en langue naturelle enregistrés en situation de travail. On remarque, en effet, que l'étude de la communication homme-machine ne peut se fonder uniquement sur le modèle de communication homme/homme même finalisé, car bien des phénomènes différents surviennent dans l'interaction homme-machine ; on se trouve ainsi en présence d'un problème circulaire [SIR 89] : celui de disposer de données pour réaliser un système qui n'a pas encore produit ces données.

Ces insuffisances des connaissances et des modèles actuels constituent l'un des obstacles majeurs des recherches visant à la conception et à la mise en œuvre d'interfaces de dialogue en langue naturelle. Les expériences de type magicien d'Oz, par exemple, où un compère humain simule les comportements de la machine à l'insu des usagers, demandent beaucoup d'effort d'organisation, pour un recueil de données relativement peu exploitable. On n'est pas sûr de surcroît, que les dialogues recueillis dans ces conditions ne sont pas quelque peu biaisés. Une autre méthode consisterait à recueillir des données issues de messageries (Minitel ou autres). On obtiendrait de cette manière des messages homme/homme, mais sans réel dialogue puisque les questions et les réponses sont trop différées dans le temps.

La méthode que nous avons choisie est donc la suivante :

- développer un système initial puis l'améliorer progressivement au vu des résultats que l'on obtient en utilisation réelle,
- en recueillant des corpus via le World Wide Web où le dialogue homme-machine est réel (et crédible) avec un temps de réponse instantané (sauf lenteur du réseau Internet).

Le problème de la première maquette se pose de manière critique. Il s'agit pour démarrer de fixer :

- (a) quel est le modèle de dialogue le plus adéquat ? en termes de stratégies de dialogue et de représentations des connaissances sur le domaine de la tâche,
- (b) quel est le système linguistique utilisé ? en termes de syntaxe et de sémantique lexicale.

Pour étudier la représentation des connaissances documentaires des utilisateurs, nous avons procédé à une simulation de la tâche de recherche d'information assistée par un vrai système de dialogue ayant une stratégie coopérative [ROU 98]. C'était une tâche simulée, car peu commune en réalité : il s'agissait, pour l'utilisateur de retrouver, grâce au dialogue, un livre que la machine avait sélectionné au hasard dans un fonds documentaire. Cela nous a permis de capturer les représentations conceptuelles du sujet et d'y associer les énoncés utilisés. Puis nous avons fait une analyse sémantique lexicale de façon à regrouper les items en champs conceptuels et à vérifier qu'ils peuvent être des déclencheurs pertinents pour la tâche. Enfin, pour ceux qui le sont et qui sont portés par des énoncés complexes, nous avons réduit l'analyse des énoncés à une analyse conceptuelle afin de modéliser le niveau de compréhension du système. Il a été possible alors d'enrichir le système de dialogue en lui ajoutant de nouvelles possibilités de compréhension et de tenir compte des stratégies observées. Le système de dialogue est soumis à épreuve auprès d'utilisateurs et les résultats sont analysés de nouveau. Le processus circulaire tourne ainsi jusqu'à convergence ou du moins jusqu'à un certain degré de stabilité et de robustesse du système.

Dans cette phase préliminaire nous avons orienté le système Halpin vers la collecte de connaissances sous forme d'un jeu interactif diffusé sur Internet. Cela nous a permis de recueillir plus de 1200 dialogues, mais seuls 980 d'entre eux ont été retenus pour le corpus Halpin-Recueil, car les autres étaient inutilisables, pour différentes raisons : problèmes de connexion au réseau Internet, trop d'insultes, etc. Nous avons effectué des analyses de dialogue et des analyses linguistiques (thématique, sémantique, syntaxique). Il en ressort principalement que :

Dans les systèmes de DHM, les premiers échanges sont importants, car l'utilisateur se crée une image mentale de son interlocuteur, et lui confère des pouvoirs et des savoirs qu'il va ensuite tenter d'utiliser, tout au long de l'interaction. Les concepts d'ouverture et de clôture de dialogue sont peu présents dans ce corpus. Ce phénomène est sans doute dû au fait que le dialogue est à l'initiative de l'usager. Le fait d'être sur le Web accentue cet effet, puisque l'on peut quitter un site aussi vite que l'on y est entré, par un simple clic souris.

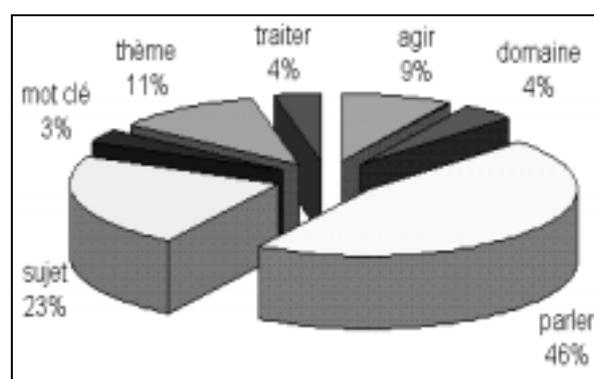


Figure 2. Le concept du « thème »

Pour obtenir des informations relatives au thème, les utilisateurs emploient souvent des phrases dans lesquelles un substantif peut être extrait : « quel est le *sujet* du livre ? », « quels sont les *mots-clés* ? », « donne moi le *thème* du bouquin », mais aussi des phrases du type « de quoi ça parle ? » où ce n'est que le verbe qui peut fournir une information sémantique (Cf. Figure 2 et Figure 3).

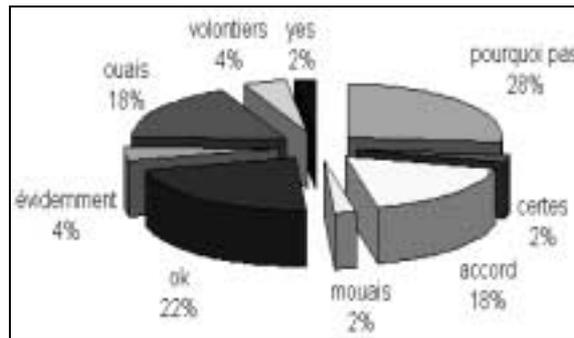


Figure 3. Le concept « acceptation » (autre que oui)

On retrouve dans 80 % des fichiers le concept « d'acceptation/confirmation ». Une très large partie des réponses sont des phrases où l'on peut déceler le mot « oui ». C'est en effet le cas pour 70 % des fichiers où l'on retrouve ce concept. Ce mot apparaît 551 fois au total dans le corpus, et 342 fois en calcul restreint². La Figure 3 montre quelques unes des autres possibilités pour ce concept. Notons par exemple que 28 % des alternatives au « oui » comportent la locution « pourquoi pas » pour exprimer une approbation. Pour le concept de « négation », le vocabulaire se limite à trois termes principaux : « non » (66 occurrences, soit 14 % de présence dans le corpus), puis « ne » et « pas » (33 et 37 occurrences, soit 7 et 8 %).

- Vocabulaire employé (verbes et substantifs) : Un autre outil d'analyse a été développé afin d'étudier la composition du corpus en terme de verbes utilisés par les usagers pour dialoguer avec la machine. Un total de 162 verbes a été identifié sur l'ensemble des fichiers du corpus. Le Tableau 1 montre la fréquence d'apparition des verbes les plus utilisés par les usagers. On note bien entendu une importante présence des verbes d'états : respectivement 24 et 88 % pour les auxiliaires « être » et « avoir » dans 117 et 425 fichiers. On retrouve également certains verbes de cognition comme chercher, rechercher, connaître, trouver, etc. A noter aussi la présence d'éléments inattendus comme « livrer » ou « titrer ». Cela s'explique vraisemblablement par une confusion (verbe/substantif) du lemmatiseur de Rank-Xerox pour les termes « livre » et « titre ».

Être	1365	Trouver	33	Chercher	16
Parler	221	Effectuer	30	Titrer	14
Avoir	177	Comprendre	25	Connaître	12
Vouloir	73	Donner	24	Livrer	12
Agir	37	Traiter	23	Appeler	11
Pouvoir	37	Rechercher	21	Poser	10
Faire	36	Aller	19	Devoir	8
Écrire	33	Dire	19		

Tableau 1. Fréquence d'apparition des verbes les plus utilisés par les utilisateurs.

Le Tableau 2 dresse une liste non exhaustive des substantifs apparus. On voit qu'ils sont majoritairement en rapport avec la tâche à accomplir. Cela servira à affiner les concepts, dans la version finale du logiciel. Par exemple, on voit bien que la seule notion de document ne suffit pas pour cette tâche ; il faudra probablement donner plus de précision à l'utilisateur : ouvrage, rapport, thèse, etc.

Livre	1081	Critère	70	Question	31
Oui	551	Sujet	66	Français	29
Auteur	411	Année	55	Parution	29
Isbn	267	Thème	51	Roman	28
Titre	266	Nom	49	Mot	27
Numéro	128	Publication	44	Domaine	23
Ouvrage	106	Langue	42	Bouquin	22
Recherche	83	Éditeur	36	Sociologie	21
Date	76	Bibliothèque	33	Chiffre	20

Tableau 2. Fréquence d'apparition des noms les plus utilisés par les utilisateurs

² termes comptabilisés une seule fois par fichier

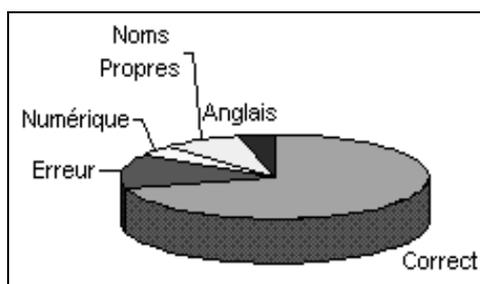


Figure 4. Vocabulaire employé dans les dialogues Homme-Machine du corpus HALPIN

Le vocabulaire global (Cf. Figure 4) utilisé par les usagers est composé de 527 termes différents : 71 % sont des mots corrects de la langue française (on trouve par exemple « chat » dans l'énoncé « je donne ma langue au chat ») ; 13 % sont des erreurs de frappe ou des mots qui n'existent pas ; 9 % sont des noms propres et 4 % des mots anglais. On retrouve également des expressions numériques dans le corpus Halpin.

En comparant ces résultats avec d'autres corpus de dialogues Homme-Machine, on note que le vocabulaire employé est souvent en étroite relation avec la tâche à réaliser, comme l'avait déjà noté Fréchet [FRE 88]. Des recherches sur la personnalisation nous ont permis de montrer de quelle manière les usagers perçoivent leur relation avec l'ordinateur [ROU 98c]. Dans notre expérience, la machine vouvoie l'utilisateur. Or, on retrouve dans les résultats autant de « tu » que de « vous ». Cependant, contrairement à d'autres recherches effectuées en protocole du Magicien d'Oz, où le « je » est marginal, ici, sa fréquence d'apparition est importante. En effet, des phrases comme « je voudrais ... », « je souhaite... », « je recherche.. » sont courantes dans le corpus Halpin.

5. Réalisation

Le système réalisé s'appuie sur un ensemble de modules implémentés sur différents serveurs distants communiquant sous le protocole IP (on utilise des logiciels du commerce, des logiciels en libre service et des logiciels développés pour Halpin).

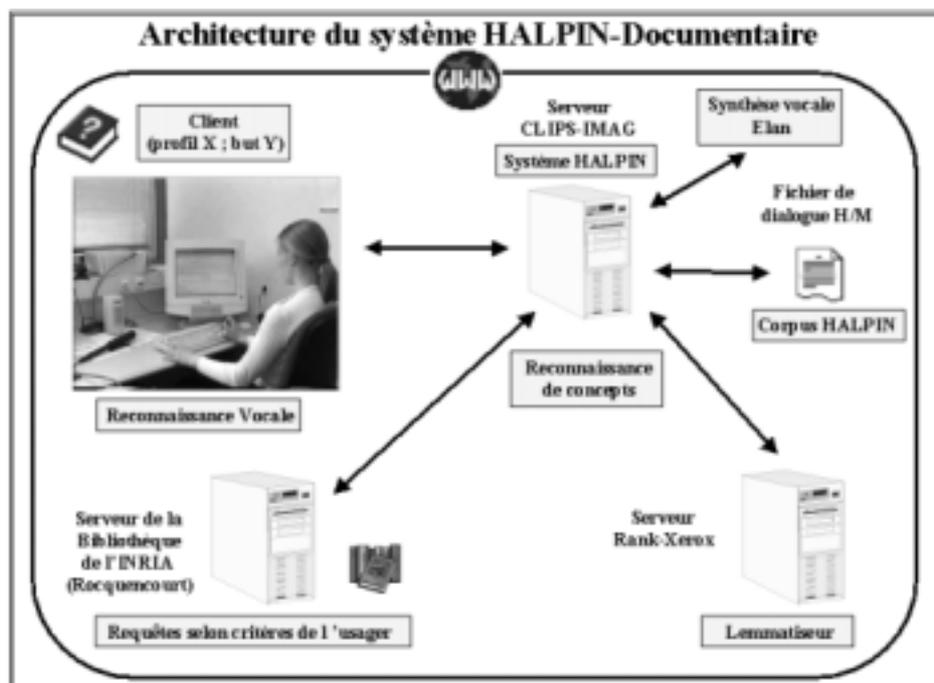


Figure 5. Architecture distribuée du système Halpin-Documentaire

L'application finale tourne dans un browser sous environnement Unix, PC, Mac, ou tout autre système supportant un navigateur gérant les applets Java. La Figure 5 présente l'architecture générale du système Halpin-documentaire.

5.1. Description des modules

- *La reconnaissance de la parole* : le système de reconnaissance vocale utilisé est celui d'IBM (ViaVoice). Ce module (sous licence) doit être installé sur la machine cliente. Il délivre en sortie une chaîne de caractères que l'utilisateur peut encore corriger à son gré avant de la valider. S'il ne dispose pas de ce logiciel, il peut entrer toutes les requêtes au clavier.

- *L'analyseur morphologique* : le lemmatiseur est celui du Centre de Recherches de Xerox [GAU 97] auquel on accède de manière distante. Ce module délivre la forme lemmatisée³ de chacun des mots de la source d'entrée.

- *La compréhension par concepts* : le module de reconnaissance de concepts a été développé dans le cadre d'Halpin [ROU 98]. Il fournit une structure de traits à partir d'une analyse conceptuelle de la source lemmatisée. Ce module est exécuté sur un serveur autonome.

- *Le contrôleur de dialogue* a été également développé dans le cadre d'Halpin [ROU 98b]. Il permet de gérer les tours de parole et l'avancée dans la tâche.

- *La génération des énoncés* de réponse est un module qui s'appuie sur un modèle de génération à "trous".

- *La synthèse vocale* est le système de la société française Elan informatique. Ce module fonctionne en serveur au laboratoire ou peut être accédé sur d'autres sites. Contrairement à la reconnaissance, l'utilisateur n'a pas besoin de l'installer sur sa propre machine.

- *L'interface graphique* et la gestion des dispositifs. Ce module tourne sous un browser comme Internet Explorer ou Netscape.

- L'accès à la base de données documentaire de l'INRIA se fait de manière distante sur le serveur dédié de l'INRIA.

Nous détaillons ci-après les modules développés dans le cadre d'Halpin : la compréhension par concepts et le contrôleur du dialogue.

5.2. La compréhension par concepts

Cette méthode de compréhension est une méthode *top-down*. Elle a l'avantage d'être robuste pour des langages finalisés contraints par la tâche car elle permet de ne pas analyser totalement l'énoncé (cela permet de tolérer les erreurs d'insertion introduites par le système de reconnaissance, et certaines erreurs de substitution ou de délétion). Cela consiste à présupposer une intention communicative et à guider l'analyse par recherche de segments-clés ou de mots-clés. Par exemple avec le but présupposé DEMANDER(DOCUMENT), l'énoncé "Donne-moi un ouvrage de Baudelaire" sera analysé comme : Donne-moi = DEMANDER, ouvrage = DOCUMENT, Baudelaire = AUTEUR, les autres mots de l'énoncé n'étant pas consommés par l'analyse. La structure de traits (attributs/valeurs) issue de l'analyse sera donc : HOMME(DEMANDER(DOCUMENT(AUTEUR = Baudelaire))). Evidemment, si l'on n'y prête pas garde l'énoncé "Donne-moi un ouvrage *sur* Baudelaire" sera analysé selon les mêmes termes au lieu de HOMME(DEMANDER(DOCUMENT(THEME = Baudelaire))). On voit donc toute l'importance de :

- Faire les bonnes hypothèses sur les buts communicationnels,
- Faire une bonne sélection des mots pertinents dans l'énoncé.

Les hypothèses sur les buts communicationnels ne relèvent pas du module de compréhension lui-même mais plutôt du module de contrôle du dialogue : c'est ce dernier qui est en effet capable de prédire les énoncés de l'utilisateur à chaque tour de parole. Par contre la sélection des mots pertinents de l'énoncé relève bien du module de compréhension. On procède pour cela en deux étapes :

1. Recherche d'expressions typées à l'aide d'un dictionnaire, pour construire des hypothèses de structures de traits

³ racine d'un mot mise dans une forme standard, par exemple *voudrais* = *vouloir+conditionnel+1èrePersSing*

2. Recherche de marqueurs grammaticaux (prépositions, négations, etc.) pouvant compléter ou orienter la structure de traits (comme l'exemple ci-dessus où *sur* oriente plutôt vers le trait *thème* que vers le trait *auteur*).

Cette méthode en deux étapes, constitue la base du processus de compréhension que nous détaillons maintenant. La Figure 6 montre les différents concepts structurés sous la forme d'un arbre de concepts.

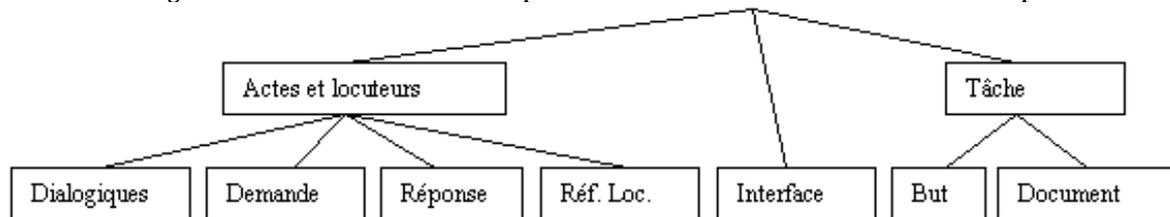


Figure 6. Arbre de concepts utilisé pour la reconnaissance de concepts dans les énoncés

- Concepts Actes et locuteurs :

Les concepts actes et locuteurs regroupent les actes dialogiques OUVERTURE ; CLOTURE ; POLITESSE ; INSULTE ; AIDÉ ; les demandes d'action RECHERCHER ; VOIR ; TOUT_VOIR ; AJOUTER ; EXECUTER ; DEMANDER ; ANNULER ; MODIFIER ; les réponses CONFIRMER ; INFIRMER ; les référents aux locuteurs MACHINE ; HOMME ; PRONOM ;

- Concepts interface :

Ces concepts se rapportent aux actions que l'on peut effectuer sur l'interface (gestion de fenêtres, de police de caractères, de périphériques vocaux) AUGMENTER ; REDUIRE ; RETABLIR ; COUPER ; sur les objets FENETRE ; POLICE ; SYNTHÈSE ; RECONNAISSANCE.

- Concepts tâche :

Les concepts tâche regroupent les buts ELARGIR ; AFFINER ; TRAITER ; EFFACER, et les référents aux documents TYPE ; NOUVEAU ; DATE ; NUMERO ; DOCUMENT ; THEME ; AUTEUR ; TITRE ; LANGUE.

Exemple de traitement du concept CONFIRMER :

Le traitement des concepts ne dépend pas des buts communicationnels. Il s'agit de faire monter d'un énoncé, tous les concepts possibles. Pour chaque concept (par exemple ici pour le concept CONFIRMER) on fait le traitement suivant :

1) recherche de mots-déclencheurs dans le sous-lexique CONFIRMER

en effet ; absolument ; acc ; accord ; admettre ; affirmatif ; amen ; assurément ; bien entendu ; bien sûr ; bien ; bon ; certain ; certainement ; certes ; confirmer ; convenir ; entendre ; évident ; évidemment ; exact ; mouais ; oc ; oil ; ouais ; oui ; oui-da ; ok ; okay ; parfait ; parfaitement ; (avec) plaisir ; positif ; positivement ; soit ; sûr ; volontiers ; voui ; tout à fait ; yep ; yes ; vrai.

2) si succès alors le concept CONFIRMER est validé,

Cette technique peut paraître assez fruste. En fait sa simplicité est un facteur de robustesse car le module de reconnaissance engendre beaucoup d'erreurs qui portent surtout sur les mots courts. Seuls les mots assez long et bien articulés ou bien accentués émergent avec une bonne probabilité. Ce sont souvent les mots les plus pertinents dans l'énoncé, puisque leur netteté est intentionnelle chez le locuteur. Un autre avantage de la méthode, réside dans la maintenance aisée des sous-lexiques qui peuvent être modifiés de manière incrémentale à l'issue d'une session de dialogue. Typiquement, si l'on observe que le mot « bouquin » à été utilisé par l'utilisateur et que la machine n'a pas su identifier le concept correspondant, on peut ajouter ce mot dans le concept « document », pour les prochaines utilisations.

Des concepts aux actes : On a donc collecté dans un énoncé tous les concepts possibles. Il s'agit maintenant de les assembler pour former une structure de traits. On utilise une grammaire sémantique à balayage gauche-droite. Sachant que les actes possibles sont les suivants : Ordonner ; Suggérer ; Confirmer ; Infirmer ; Demander ; Aider ; Saluer ; Insulter ; Remercier ; Alerter ; Justifier ; Commander ; Promettre ; Informer ;

Ainsi l'énoncé "oui d'accord, merci, je souhaite rechercher par nom d'auteur" donnera la suite d'actes structurés suivants :

HOMME(CONFIRMER)+HOMME(REMERCIER)+HOMME(DEMANDER(CRITERE=AUTEUR))

Ces trois actes seront exécutés en séquence dans le cours du dialogue.

On remarquera que cette technique permet de traiter les auto-corrrections "oui, d'accord, attendez... non" par HOMME(CONFIRMER)+HOMME(INFIRMER), mais ne permet pas de traiter les ambiguïtés ou les contradictions comme "oui et non", "oui, mais je ne suis pas d'accord".

5.3. Le contrôleur de dialogue

Le modèle de dialogue que nous avons choisi pour Halpin, est un modèle de dialogue à stratégie dynamique, pour apporter à la recherche d'informations les avantages décrits au §3.

Le modèle utilise trois stratégies :

- Directive pour les phases d'ouverture, d'introduction et de clôture (c'est-à-dire à initiative de la machine). Pendant la phase d'ouverture on tente de connaître le but de l'utilisateur et on lui demande le niveau d'assistance qu'il préfère (au moins au départ). C'est une phase importante pour la réussite de la suite du dialogue.

- Réactive pour les usagers qui le souhaitent. On fait ensuite l'hypothèse que ces utilisateurs savent ce qu'ils cherchent et comment l'obtenir. Il n'y a donc pas lieu d'alourdir le dialogue de questions qui leur paraîtraient saugrenues.

- Coopérative pour les usagers qui le souhaitent. On suppose ici que ces usagers ne maîtrisent pas leur environnement de recherche soit au niveau de la tâche elle-même, soit au niveau de l'interface.

La stratégie directive est guidée par les buts de la machine : connaître le but de l'utilisateur et son niveau d'assistance. La stratégie coopérative est guidée par les buts de l'utilisateur que la machine tente d'inférer des actes de dialogue, tandis que la stratégie réactive est dirigée par les données, c'est-à-dire ici par les énoncés de l'utilisateur, autrement dit par les concepts identifiés par la machine.

Ces stratégies peuvent être rapprochées de la notion de contrôle du dialogue, définie et utilisée dans les systèmes de gestion d'Interface utilisateur [GAI 81], dans lesquels le contrôle du dialogue précise qui a l'initiative du dialogue. Il peut être interne (à l'initiative du système), externe (à l'initiative de l'utilisateur), ou mixte (à l'initiative partagée).

Pour la mise en œuvre du contrôleur de dialogue, nous nous sommes inspirés du modèle COR (Conversational Roles) de [STE 95]. *As information seeking and retrieval are interactive processes, we believe that providing a flexible and co-operative human-machine dialogue is a complementary way to improve information retrieval systems* [BAT 95]. Ce modèle propose des séquences de dialogues *idéales* et *alternatives*. Par exemple, un dialogue entre A (qui demande de l'information) et B (qui en fournit) est formalisé comme suit :

Demande (A,B) => promesse (B, A) + information (B,A) + test-satisfaction (A,B)

Ce modèle peut être amélioré en intégrant au sein du processus interactif des données relatives à l'utilisateur et à ses stratégies. Notre modèle est dérivé de celui de Stein [STE 97] mais augmenté par la prise en compte des buts et du niveau de compétences de l'usager, de sorte que le système puisse adapter ses réponses au niveau d'assistance utilisateur et à la tâche en cours. Le modèle que nous proposons pour la demande, est le suivant (il en serait une formulation similaire pour les autres actes de parole) :

HOMME(DEMANDER)/[NIVEAU D'ASSISTANCE(HOMME)+BUT(HOMME)] => MACHINE (PROMETTRE) + MACHINE(INFORMER) + MACHINE{JUSTIFIER + SUGGERER}

Les notations [] signifiant en "contexte de" et {} présence facultative. Dans l'algorithme ci-dessous HOMME sera abrégé en H et MACHINE en M. Il est évident que H s'adresse à M et que M répond à H, nous économisons donc la notation du sens A,B de Stein, ci-dessus.

Ce qui se lit : la réponse à une demande de HOMME par MACHINE est conditionnée par le niveau d'assistance de HOMME et son but. Cette réponse sera constituée d'une promesse (mise en attente), d'une information sur l'objet de la demande et éventuellement suivie d'une justification et d'une suggestion coopérative pour le tour de parole suivant. Le modèle de dialogue utilisé adopte donc des stratégies différentes en fonction du niveau d'assistance de l'utilisateur et du but. Ci-dessous, un extrait de l'algorithme implémenté montre comment la machine gère le dialogue en fonction des concepts détectés, des buts ainsi que des actes de dialogue :

```

Début-Dialogue
Début Ouverture
Affichage de la page d'accueil Web
Stratégie = Directive
Sous-dialogue But (B)
Sous-dialogue User (Assistance)
Cas Assistance == aidée ; Stratégie = Coopérative ; M(Justifier(stratégie))+M(Promettre(aide))
Cas Assistance == non-aidée ; Stratégie = Réactive ; M(Justifier(stratégie))
Fin Ouverture
A = H(Début) ; Tour Parole = 0 ; Item = Thème
TantQue A ≠ H(Quitter) Faire :
Cas Assistance == aidée ; M(Demander(Item)) ; Attente(A) ; Compréhension(A)
    Cas A == H(Informer) ; si Item == vide alors Sous-dialogue Clarification(Item)
    Cas A == H(Ordonner) ; si Item == vide alors M(Alerter), sinon
M(Promettre)+M(Informer(N))+M(Suggérer(Item suivant))
    Cas A == H(Confirmer) ; M(Relancer(Demander(Item)))
    Cas A == H(Informer) ; Item = Item suivant ; M(demander(Item))
    Cas A == H(Demander(aide)) ; M(Informer(aide))+M(Demander(changer(Assistance)))
    Cas A == H(Saluer) ; Si Tour Parole>3 alors clôture sinon M(Saluer)
    Cas A == H(Insulter) ; Si Compteur insulte < 2 alors M(Menacer), sinon M(Insulter)
    Cas A == H(Remercier) ; M(Remercier)+M(Relancer(Demander(Item)))
    Cas A == H(Commander) ; M(Exécuter(commande-interface))
Item = Item suivant ; Tour Parole = Tour Parole + 1
Cas Assistance == non-aidée : Attente(A) ; Compréhension(A) ; Analyse(A, Item)
    Cas Item == Interface ; M(Exécuter(commande-interface))
    Cas Item == Auteur ; Remplir Critère(Auteur) ; M(Justifier)
    Cas Item == Date ; Remplir Critère(Date) ; M(Justifier)
    Cas Item == Type_document ; Remplir Critère(Type_document) ; M(Justifier)
    Cas Item == Titre ; Remplir Critère(Titre) ; M(Justifier)
    Cas Item == Inconnu ; M(Justifier)+M(Relancer(Demander(Item)))
    Cas Item == Aide ; M(Aider)+M(Suggérer(Changer(Assistance)))
    Cas A == H(Informer) ; si Item == vide alors Sous-dialogue Clarification
    Cas A == H(Ordonner) ; si Item == vide alors M(Alerter), sinon M(Promettre)+M(Informer(N))
    Cas A == H(Confirmer) ; M(Suggérer)
    Cas A == H(Informer) ; Item = Item suivant ; M(demander(Item))
    Cas A == H(Demander(aide)) ; M(Informer(aide))+M(Suggérer(Changer(Assistance)))
    Cas A == H(Saluer) ; Si Tour Parole>3 alors clôture sinon M(Saluer)
    Cas A == H(Insulter) ; Si Compteur insulte < 2 alors M(Menacer), sinon M(Insulter)
    Cas A == H(Remercier) ; M(Remercier)
    Cas A == H(Commander) ; M(Exécuter(commande-interface))
    Tour Parole = Tour Parole + 1
FinTantQue
Clôture ; M(Demander(Remplir(Questionnaire)))
FinDialogue

Sous-dialogue Clarification (Item)
    B = Début
    TantQue B ≠ Quitter Faire :
M(Demander(B)) ; Attente(B) ; Compréhension(B) ; Tour Parole = Tour Parole + 1
    Cas B == Elargissement ; M(Supprimer(Item))+M(Promettre)+M(Informer(N))
    Cas B == Inconnu ; M(Demander(satisfaction))
    Cas B == Affinement ; M(Demander(Item))
    Cas B == Quitter ; FinSous-dialogue
    FinTantQue
FinSous-dialogue

```

Où N est le nombre de réponses du système (nombre de documents satisfaisant la requête).

NB : nous ne décrivons pas ici les autres sous-dialogues.

6. Résultats

Les résultats que nous exposons ici sont plus d'ordre qualitatif que quantitatif (une évaluation a été faite sur un panel de sujets en laboratoire). Nous exposons tout d'abord des résultats assez généraux, puis nous présentons les tendances qui se dégagent quant à la satisfaction des usagers lors de l'utilisation d'Halpin sur le Web.

6.1. Résultats généraux

Les résultats montrent que les utilisateurs dialoguent volontiers avec la machine grâce notamment au module intégré de synthèse vocale⁴ [ELA 01] qui est un fort incitateur de maintien du dialogue (24 sujets ont été analysés sur des tâches de difficulté croissante : recherche d'un document dans la base avec des spécifications de plus en plus imprécises). Les avantages de l'utilisation du langage naturel sont aussi démontrés : utiliser un vocabulaire non restreint, nommer des objets (ou des concepts) qui ne sont pas forcément présents à l'écran, les nommer de manière imprécise. Le dialogue en langage naturel arrive majoritairement dans l'ensemble des tâches et complète le mode de navigation : toute chose étant égale par ailleurs (sur une base de données identique), une interaction multimodale est une réponse efficace aux problèmes de désorientation, de surcharge cognitive, et d'évaluation de la pertinence des réponses du système. Les exemples de dialogues réels, ci-dessous, recueillis avec le système Halpin montrent comment la machine se comporte dans le dialogue.

11:59:18:M0=Bonjour. Que souhaitez-vous faire ? 1 : Retrouver un document que vous connaissez, 2 : Rechercher un ensemble de documents ou 3 : Découvrir ce site ?
12:00:03:H1=2
12:00:15:M2=Rechercher un ensemble de documents. D'accord. Souhaitez-vous que je vous aide à formuler votre requête ?
12:00:30:H2=oui
12:00:37:M3=Pouvez-vous me donner le thème recherché, entre guillemets s'il vous plaît ?
12:01:11:H3=karaoke
12:01:28:M4=Vous pouvez soit me donner le thème entre guillemets, soit effacer le thème en cours. Ou bien dites Non pour passer à autre chose.
12:02:05:H4=je cherche des documents sur le "karaoke "
12:02:18:M5=Je n'ai pas trouvé de document ayant comme thème karaoke . Voulez-vous modifier votre requête ?
12:03:51:H7=je veux chercher avec "synthese" please
12:04:09:M8=J'ai trouvé 398 documents ayant comme thème synthese . Voulez-vous affiner votre requête ?
12:05:28:H8=je veux l'annee 1998
12:06:01:M10=J'ai trouvé 12 documents de 1998 ayant comme thème synthese . Voulez-vous affiner votre requête ?
12:06:46:H10=je ne vois rien
12:06:58:M11=Okay j'augmente la taille de la police.
12:07:09:H11=efface tout
12:07:24:M12=D'accord, tous les critères sont vides à présent. Faites un choix parmi les critères : Auteur, Titre, Thème, Année, ou Type de document.
12:07:35:H12=auteur
12:07:49:M13=Vous pouvez soit m'indiquer le nom de l'auteur, avec une majuscule, soit effacer l'auteur en cours. Ou bien dites Non pour passer à autre chose.
12:08:08:H13=je veux Astrom
12:08:22:M14=J'ai trouvé 5 documents avec comme auteur Astrom. Consultez-les sur la page de droite. Est-ce que cela vous convient ?
12:10:20:H17=voui
12:10:30:M18=Très bien, je suis heureux d'avoir pu vous aider dans votre recherche.

Capture 1. Extrait d'un dialogue du corpus Halpin-Documentaire

Dans cet exemple (Capture 1), on observe bien entendu des tours de parole classiques à propos de la recherche d'informations, mais on en note également d'autres concernant l'interface, comme par exemple l'intervention *H10=je ne vois rien*, à laquelle la machine répond par *M11=Okay j'augmente la taille de la police*.

(...)
15:53:17:M=Rechercher un ensemble de documents. D'accord. Souhaitez-vous que je vous aide à formuler votre requête ?
15:53:43:H=ok

⁴ Speech Engine de Elan Informatique.

15:53:52:M=Pouvez-vous me donner le thème recherché, entre guillemets s'il vous plaît ?
 15:57:02:H=je souhaite des references en intelligence artificielle
 15:57:15:M=Vous pouvez soit me donner le thème entre guillemets, soit effacer le thème en cours. Ou bien dites Non pour passer à autre chose.
 15:57:35:H=je veux des trucs en "intelligence artificielle"
 15:57:58:M=J'ai trouvé 1204 documents ayant comme thème intelligence artificielle . Voulez-vous raffiner votre requête ?
 15:58:18:H=oui, il y en a beaucoup trop mon pauvre
 15:59:07:M=Le thème actuel est intelligence artificielle . On pourrait y ajouter par exemple : representation connaissance ; langage naturel ; resolution probleme ; apprentissage ; acquisition connaissance. Pour ajouter un thème, donnez le entre guillemets, sinon répondez Non pour passer à autre chose.
 16:00:21:H=oui, voyons avec "langage naturel"
 16:00:51:M=J'ai trouvé 174 documents ayant comme thème intelligence artificielle langage naturel . Voulez-vous raffiner votre requête ?
 16:01:04:H=ben oui
 16:01:23:M=Afin de raffiner, vous devez choisir parmi les critères : Auteur, Titre, Thème, Année, ou Type de document.
 16:01:58:H=l'annee 98
 16:02:17:M=J'ai trouvé 4 documents de 1998 ayant comme thème intelligence artificielle langage naturel. Consultez-les sur la page de droite. Est-ce que cela vous convient ?
 16:02:42:H=ouais, ca a l'air bon
 16:02:51:M=Très bien, je suis heureux d'avoir pu vous aider dans votre recherche.

Capture 2. De 1204 à 4 réponses : la machine propose des thèmes proches de celui donné.

Avec l'exemple de la Capture 2, on a typiquement un cas où le système Halpin aide l'utilisateur de manière coopérative, en tenant compte de son but. On passe donc de 1204 réponses à 4 réponses par raffinement de la recherche. Pour cela, il lui propose des termes proches du thème initialement demandé, se focalise sur l'année voulue (il comprend que 98 signifie 1998) et s'informe de la satisfaction de l'utilisateur par rapport aux documents qu'il lui propose. Dans d'autres situations, il permet de retrouver l'orthographe correcte d'un auteur ou encore de n'obtenir que les documents les plus récents (sans préciser d'années), choses que les moteurs actuels de recherche d'informations ne font pas.

6.2. Évaluation du système Halpin

Les usagers qui se sont connectés librement au système Halpin avaient la possibilité de répondre à un questionnaire d'évaluation facultatif, en ligne. Nous présentons ci-dessous quelques un des résultats obtenus, bien que l'échantillon des personnes ayant répondu à ce questionnaire de satisfaction soit encore faible, par rapport au nombre de connexions effectivement enregistrées. *Globalement, l'interaction dialoguée avec le système HALPIN est considérée, par les 19 personnes interrogées, comme bonne à 56 %, mauvaise à 33 % et très mauvaise à 11 %. A la question « Le dialogue avec la machine vous a-t-il orienté dans votre recherche ? », 24 % ont répondu « Pas assez », 47 % « Suffisamment », et 29 % « Un peu ».*

Les usagers déclarent à 59 % que le dialogue leur a permis de trouver ce qu'ils cherchaient (*en partie*: 23,53 % ; *un peu* : 17,65 % ; *totalemment* : 17,65 %) contre 41 % n'ayant *pas du tout* retrouvé l'information recherchée. Le dialogue est perçu comme robuste pour la moitié des personnes interrogées, et 56 % ont apprécié l'aide apportée par la machine lors des interactions dialoguées.

La synthèse vocale utilisée est considérée comme « *très bonne* » pour 21 % des utilisateurs, « *bonne* » pour 47 % et « *mauvaise* » pour 10 % (21 % de sans opinion). Elle leur semble très intelligible (42 %) et moyennement naturelle (37 %). Les avis à propos de l'utilité du système, comparé à un formulaire traditionnel du Web, sont présentés sur la Figure 7 ci-dessous. Dans l'ensemble, 41 % des personnes trouvent ce système moins utile qu'une interaction traditionnelle, 12 % la trouvent similaire et 47 % la trouvent meilleure.

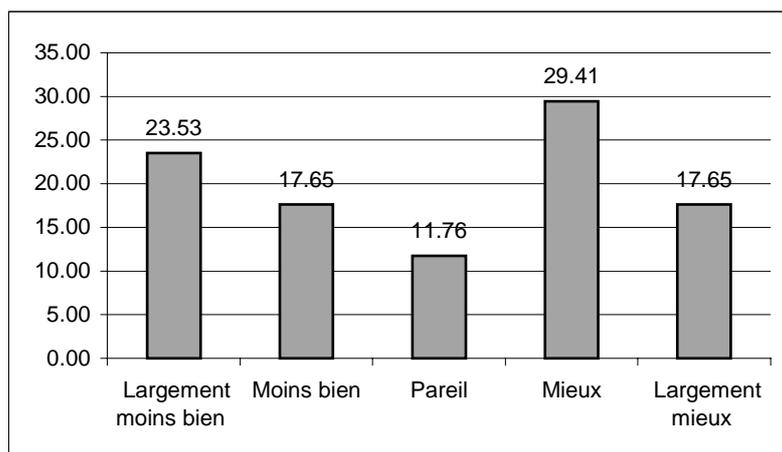


Figure 7. Utilité du système par rapport à un formulaire traditionnel du Web

Les temps de réponse du système, quant à eux, sont jugés comme étant *rapides* à 16 %, *acceptables* à 37 % et *trop longs* à 47 %. La majorité des personnes ayant répondu sont des hommes (79 %). La plupart sont des informaticiens (58 %) (et 12 % de bibliothécaire). Notons enfin que les salariés sont représentés à 42 %, les enseignant et/ou chercheurs à 32 %, et les étudiants à 26 %.

7. Conclusion

Ces travaux valident nos hypothèses de base, à savoir qu'un système de dialogue multimodal est efficace en recherche d'information. Il permet d'une part d'adapter sa stratégie par rapport à l'avancée de la tâche et d'autre part de profiter de la souplesse d'expression qu'offre la langue naturelle. L'aspect sémantique devient alors une composante forte pour un dialogue coopératif et finalisé : nous avons montré qu'un modèle de DHM basé sur une reconnaissance de concepts est viable pour cette tâche, car elle est fortement structurée par les buts. L'adaptation de la machine à son interlocuteur (niveau d'assistance, but, etc.) permet de rendre le DHM plus souple et plus performant : on évite les aides et une assistance prégnante pour les experts par exemple, on assiste davantage les novices et on relâche cette assistance à la demande. Plus encore, les utilisateurs apprécient que la machine leur parle et qu'ils puissent lui répondre de façon spontanée et naturelle, via le Web, malgré des temps de réponse encore un peu longs.

Un des intérêts du système réside aussi dans l'intégration de ressources déjà existantes (base de données de l'INRIA, outil morphologique de Xerox, reconnaissance de la parole d'IBM, synthèse de la parole d'Elan Informatique) qui montre la possibilité de ré-utiliser certains composants dans une architecture ouverte. En contre partie, le système demeure fragile du fait même de sa dépendance à d'autres serveurs : il se trouve en défaut si une seule de ces ressources n'est plus disponible. Des améliorations devront être apportées quant à la robustesse du système lors d'erreurs et de mauvaises interprétations des énoncés par le système de reconnaissance vocale. Une meilleure gestion des procédés anaphoriques⁵, tout comme les ellipses⁶ ou les tropes⁷ est à prendre en compte, car ils assurent la concision et la souplesse d'un dialogue en LN. Pour aller plus loin, nous devons fournir la possibilité d'utiliser une reconnaissance vocale distante, et non plus seulement locale, comme c'est le cas actuellement avec IBM ViaVoice. Le système RAPHAEL [AKB 98], [SCH 97] est actuellement en cours d'intégration dans le système Halpin est sera disponible pour la prochaine version. Cela devrait permettre un accès au système par des postes clients ne possédant ni logiciel de synthèse, ni de reconnaissance vocale sur leur proche machine ; tout cela étant pris en charge par le serveur du système Halpin. Les problèmes d'attente (entraînant une rupture du caractère naturel du dialogue) que dénoncent principalement les utilisateurs devraient donc pouvoir être surmontés dans quelques temps. Mais plus que sur les problèmes d'ordre technique, c'est en matière de robustesse et de pertinence du modèle de DHM que nous allons travailler : en intégrant un puissant thesaurus français [MEM 01], nous pensons pouvoir sensiblement améliorer, d'une part la compréhension des entrées, et d'autre part la diversité des répliques, de manière à assurer un dialogue ouvert, consistant et le plus naturel possible, en dépassant le cadre de la recherche

⁵ L'anaphore est un mot renvoyant à un membre de phrase déjà énoncé.

⁶ L'ellipse est l'omission d'un ou de plusieurs mots d'un énoncé sans altération du sens de l'énoncé.

⁷ Un trope est une figure rhétorique consistant à employer un mot ou une expression en les détournant de leur sens propre.

documentaire pour l'étendre à la recherche d'information sur le Web (passage de documents structurés aux documents non structurés).

8. Bibliographie

- [AKB 98] AKBAR M., CAELEN J., Parole et traduction automatique : le module de reconnaissance RAPHAEL, *COLLING-ACL'98*, pp. 36-40, Montréal (Québec), Août 1998.
- [BAL 96] BALPE, J.-P., LELU, A., PAPY, F. & SALEH, I. *Techniques avancées pour l'hypertexte*, Paris: Ed. Hermès, 1996.
- [BAT 95] BATEMAN, J. A.; HAGEN, E., & STEIN, A. Dialogue modeling for speech generation in multimodal information systems, in P. Dalsgaard, et al. (Ed.), *Proceedings of the ESCA Workshop on Spoken Dialogue Systems - Theories and Applications* (pp. 225-228). Aalborg, Denmark: ESCA/Aalborg University, 1995.
- [BEL 97] BELISLE, C., ZEILIGER, R., CERRATTO, T., (1997), Integrated Cognitive Engineering at the Interface, in *proceedings of the Second International Conference on Cognitive Technology, CT'97*, Aizu, Japan, Marsh, Hehaniv, Gorayska eds., IEEE Computer Society, Los Alamitos.
- [BUC 98] BUCH, P., A l'heure d'Internet, *Cahiers Pédagogiques, N°362*, Mars 1998.
- [CAR 96] CARD, S.K., ROBERTSON, G., G., YORK, W., The WebBook and the Web Forager : An Information Workspace for the World-Wide-Web, in *Proceedings of the CHI'96 conference, ACM press*, 1996.
- [CAS 96] CASTELLI, C. COLASSO, L. & MOLINARI, A., Getting lost in Hyperspace: Lessons Learned and Future Directions, in *ED-MEDIA 96/ED-TELECOM 96*, 1996.
- [CON 87] CONKLIN, J., Hypertext: an introduction and survey, *IEEE Computer*, pp. 17-41, September 1987.
- [ELA 01] <http://www.elan.fr>
- [FRE 88] FRECHET, A-L, Analyse linguistique d'un corpus de dialogue Homme-Machine (oral finalisé), *Thèse, Université de la Sorbonne nouvelle, Paris III*, 1988.
- [GAI 81] GAINES, B.R., The Technology of Interaction - Dialogue Programming Rules. *International Journal of Man-Machine Studies* Vol. 14, pp 133 - 150. Academic Press, Londres, 1981.
- [GAU 97] GAUSSIÉ, E., GREFFENSTETTE, G., SCHULZE, M., Traitement du langage naturel et recherche d'informations : quelques expériences sur le français. *Premières Journées Scientifiques et Techniques du Réseau Francophone de l'Ingénierie de la Langue de l'AUPELF-UREF*, Avignon, Avril 1997.
- [HAR 98] HARPER, R., Information that Counts : Socioly, Ethnography and Work at the International Monetary Fund, in *Proceedings of Workshop on Personalised and Social Navigation in Information Space, Hook, Munro, Benyon eds., SICS Technical Report T98: 02*, Kista, Sweden, 1998.
- [MEI 98] MEIYE, J.P., ROUILLARD, J., VAUFREDAZ, D., Webcompletion - protocole de normes associatives sur Internet. *École thématique CNRS - Sémantique*, Caen (Asnelles-sur-mer), Janvier 98.
- [MEM 01] <http://www.memodata.com/>
- [ORI 99] <http://www.gate.cnrs.fr/~zeiliger/Orion99.doc>
- [RAB 95] RABARDEL, P., Les Hommes et les Technologies : approche cognitive des instruments contemporains. Paris: A.Colin, 1995.
- [ROU 97] ROUET, J-F., Le lecteur face à l'hypertexte, in Grimont, Alain, sous la direction de, *Apprendre avec le multimédia Où en est-on?*, Ouvrage collectif, Paris: Ed. Retz, 1997.
- [ROU 98] ROUILLARD, J. et CAELEN, J., Étude 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.
- [ROU 98a] ROUILLARD, J., Les enjeux d'un dialogue Homme-Machine sur Internet - L'Hyperdialogue. *Bulletin d'informatique approfondie et applications, revue de l'université de Provence*, n° 52, Mars 99.
- [ROU 98b] ROUILLARD, J., Hyperdialogue Homme-Machine sur le World Wide Web : Le système HALPIN, *Colloque International Ergonomie et Informatique, ERGO'IA*, Biarritz, Novembre 98.
- [ROU 98c] ROUILLARD, J., Contribution à l'étude du dialogue Homme-Machine à travers le Web : la personnalisation, *RECITAL'98*, Le Mans, Septembre 98.
- [ROU 00] ROUILLARD, J., Hyperdialogue sur Internet. Le système HALPIN, Thèse de doctorat d'informatique, Université Grenoble I, 2000.
- [SCH 97] Schultz T., Westphal M., Waibel A., The GlobalPhone Project: Multilingual LVCSR with JANUS-3, Multilingual Information Retrieval Dialogs: 2nd SQEL Workshop, pp 20--27, Plzen, Czech Republic, April 1997.
- [SIR 89] SIROUX, J., GILLOUX, M., GUYOMARD, M., SORIN, C., *Le dialogue homme-machine en langue naturelle : un défi ?* Annales Télécommunication, 44, n° 1-2, 1989.
- [SMA 98] SMAÏL, M., « Vers des systèmes évolutifs de recherche d'information : un état de l'art ». *Revue TSI : Technique et Science Informatique, Vol. 17, n°10*, décembre 1998.
- [STE 95] STEIN, A. & MAIER, E., Structuring collaborative information-seeking dialogues, *Knowledge-Based Systems, 8(2-3, Special Issue on Human-Computer Collaboration): 82-93*, 1995.

[STE 97] STEIN, A., GULLA, J. A., MÜLLER, A. & THIEL, U., Conversational interaction for semantic access to multimedia information, in *M.T. Maybury (Ed.), Intelligent Multimedia Information Retrieval (pp. 399-421). Menlo Park, CA: AAAI/The MIT Press, 1997.*

[SOU 96] SOUZA, A.,P., DIAS, P., Analysis of Hypermedia browsing processes in Order to Reduce Disorientation, in *Proceedings of ED-MEDIA '96 conference, AACE, 1996.*

9. Remerciements

Les auteurs tiennent à remercier la Région Rhône-Alpes, qui leur a permis d'effectuer ce travail ; la bibliothèque de l'INRIA ; et l'équipe Xerox de Meylan (Eric Gaussier, Grégory Grefenstette, et Maximilian Schulze) qui les a autorisés à utiliser l'outil de lemmatisation disponible sur le Web ; sans oublier les membres de l'équipe GEOD qui ont travaillé sur le traitement du signal et sur des problèmes réseaux : Mohammad Akbar, Harouna Kabré et Dominique Vaufreydaz.