



HAL
open science

Hyperdialogue sur internet. Le système HALPIN

José Rouillard

► **To cite this version:**

José Rouillard. Hyperdialogue sur internet. Le système HALPIN. Interface homme-machine [cs.HC].
Université Joseph-Fourier - Grenoble I, 2000. Français. NNT : . tel-00006753

HAL Id: tel-00006753

<https://theses.hal.science/tel-00006753>

Submitted on 25 Aug 2004

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

UNIVERSITE JOSEPH FOURIER – GRENOBLE 1
SCIENCES & GEOGRAPHIE

THESE

pour obtenir le grade de

DOCTEUR DE L'UNIVERSITE JOSEPH FOURIER

Discipline : Informatique

Présentée et soutenue publiquement

par

José Rouillard

Le 11 janvier 2000

HYPERDIALOGUE SUR INTERNET

Le système HALPIN

COMPOSITION DU JURY :

Marie-France Bruandet,	Examineur
Jean Caelen,	Directeur de thèse
Daniel Luzzati,	Examineur
Violaine Prince,	Rapporteur
Jacques Siroux,	Rapporteur

Thèse préparée au sein du laboratoire de Communication Langagière et Interaction

Personne-Système - CLIPS - Fédération IMAG

Remerciements

Tout d'abord, je tiens à remercier mon directeur de thèse, *Jean Caelen*, de m'avoir accueilli dans son équipe, et de m'avoir accordé sa confiance tout au long de ce travail.

Je remercie chaleureusement les membres de mon jury, *Marie-France Bruandet*, *Daniel Luzzati*, *Violaine Prince* et *Jacques Siroux* pour l'honneur qu'ils m'ont fait en acceptant d'examiner ce travail.

Je tiens également à remercier *Romain Zeiliger* et *André Bisseret*, membres du projet Orion, pour tout l'intérêt qu'ils ont porté à notre collaboration, et pour leurs appréciables conseils lors de nos fructueuses discussions. Je n'oublie pas la *Région Rhône-Alpes*, pour son soutien scientifique et financier.

Je pense également à mes fidèles amis, *Roger Desfournaux*, *Christophe « Dataman » Hardy*, *Pierre-Luc Arnaud*, ainsi qu'aux personnes qui ont cru en moi et qui ont su m'éclairer lorsqu'il fallait faire des choix : *Yvan Raïtchef*, *Pierre Lagrue*, *Jean-Paul Arnaud*, *François Hoynant*, *Jacques Broda*, *Françoise Adreit*, *Jean-Michel Knippel*, *Edmond Bianco*.

Un grand merci à tout le personnel du laboratoire CLIPS, et aux personnes qui m'ont permis de réaliser cette thèse, avec une mention spéciale pour ces excellents programmeurs que sont *Dominique Vaufreydaz* et *Mohammad Akbar*.

Je souhaite encore remercier, *Éric Castelli*, *Solange Hollard*, *Harouna Kabré*, et surtout, *Dominique Lalane*, mes précieux relecteurs, pour avoir effectué cette tâche difficile, avec attention et minutie.

Je remercie également deux personnes qui me sont chères, ma *Maman* et *Domie*, pour leur soutien, leurs remarques et leur confiance.

Que toutes ces personnes reçoivent ici ma profonde gratitude.

Grenoble, le 26.01.2000

À ma mère.

SOMMAIRE

1	INTRODUCTION	6
2	SITUATION DU PROBLÈME : DIALOGUE ET RECHERCHE D'INFORMATION	16
3	BESOINS DES UTILISATEURS EN SITUATION DE RECHERCHE D'INFORMATION	43
4	MISE EN ŒUVRE D'UN SYSTÈME DE DHM	53
5	LE SYSTÈME HALPIN-RECUEIL	61
6	LE SYSTÈME HALPIN-DOCUMENTAIRE	99
7	RÉSULTATS ET ÉVALUATIONS DU SYSTÈME HALPIN-DOCUMENTAIRE	141
8	CONCLUSION ET PERSPECTIVES	163
	BIBLIOGRAPHIE	169
	TABLE DES ILLUSTRATIONS	181
	ANNEXES	184

1 Introduction

*"Notre civilisation, dis-je, croit continuer les autres parce qu'elle leur succède.
Or elle rompt avec elles par un fait banal et capital, la machine."*

André Malraux, « Le miroir des limbes ».

Cette thèse a pour objet le dialogue écrit et oral, entre un interlocuteur humain et une machine, via le réseau Internet. Nous faisons d'abord quelques rappels concernant le dialogue homme-machine et le réseau Internet, puis nous présentons notre méthodologie de travail, fondée, pour l'essentiel, sur l'usage.

1.1 Le Dialogue Homme-Machine

Le traitement automatique de la langue naturelle (TALN), et le dialogue homme-machine (DHM) sont des disciplines qui suscitent un intérêt particulier auprès des chercheurs en informatique, en linguistique, et en sciences cognitives. 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 conçus dans cet objectif, avec plus ou moins de réussite.

Même si communiquer avec autrui s'avère un mécanisme complexe, quels que soient la situation et le contexte, un dialogue avec une machine présente un intérêt certain : « *Exprimer son problème en langage naturel à une machine, n'est-ce pas le but ultime des recherches en programmation ?* » [COULON et KAYSER 86].

Le DHM est, selon nous, une forme particulière de communication avec une machine. La plupart du temps, nous communiquons avec nos ordinateurs, sans pour autant dialoguer avec eux. Les machines répondent à des commandes, et exécutent les ordres demandés. Elles ne sont pas programmées pour « comprendre » ce qu'elles font. *« Lorsque l'on utilise un traitement de texte par exemple, le clavier assure le canal de communication homme-machine, et l'écran celui de la communication machine-homme, et l'on ne peut pas pour autant dire qu'il s'agit là véritablement de dialogue. »* [LUZZATI 95].

Certes, une grande partie de nos interactions avec les machines ne requièrent aucune forme « d'intelligence » de leur part. Il existe cependant des circonstances dans lesquelles une communication langagière peut s'avérer fructueuse. *« Nous ne prétendons pas que toute communication homme-machine doit systématiquement utiliser une langue naturelle, mais il est des situations où celle-ci est indéniablement utile : en particulier quand l'utilisateur est incapable d'explicitier précisément ses besoins, quand il refuse d'apprendre un code rébarbatif, quand la tâche n'est pas bien spécifiée, ou encore avec des interfaces non conviviales. »* [SABAH et al. 97].

Nos travaux n'ont pas pour objectif de donner aux machines la possibilité de tenir toutes sortes de conversations avec un interlocuteur humain. Weizenbaum a montré, grâce à son système Éliza [WEIZENBAUM 66], qu'un dialogue non finalisé, avec une machine, est tout à fait réalisable. Mais ce genre de simulations atteint ses limites lorsque l'utilisateur souhaite obtenir des informations auprès de la machine, sur la base d'un dialogue constructif, en vue de réaliser une tâche déterminée, et qu'elle ne saisit manifestement pas ce qu'il lui demande.

Bien au contraire, nous souhaitons mettre en œuvre un système qui analyse et interprète des énoncés en langue naturelle (LN), afin d'aider l'utilisateur à atteindre un but. On fait ainsi référence à un dialogue finalisé : *« De toute façon, on ne sait vraiment réaliser de tels modules de gestion que dans le cadre de dialogues finalisés [PIERREL 87], où l'influence de la tâche a une importance essentielle. »* [SABAH 89].

Ce qui amène la constatation suivante : « *La question qui se pose dans le dialogue homme-machine peut être résumée comme la recherche des processus qui interviennent dans la compréhension et la production des échanges dialogiques dans une situation déterminée.* » [VILLASEÑOR-PINEDA 97].

Si cette approche demeure délicate, cela est essentiellement dû aux difficultés rencontrées pour « faire entrer la langue dans la machine » [COURSIL 92], [RASTIER 87]. Les chercheurs en TALN travaillent depuis longtemps dans cette voie. L'exercice est d'autant plus périlleux lorsque l'on prend conscience, que paradoxalement, le dialogue est à la fois « *une des capacités qui nous est les plus familières* » et « *quelque chose que l'on ne maîtrise pas* » vraiment (au niveau lexical, syntaxique, sémantique et pragmatique) [BILANGE 92], [CARRÉ et al. 91].

Nos travaux de thèse tentent de répondre à certaines interrogations, à propos du DHM, que Sabah résume en ces mots : « *Une autre question à laquelle il conviendrait de répondre est : quel intérêt réel présentent les phénomènes traités pour le dialogue homme-machine ? En d'autres termes, dans une vraie situation de dialogue avec un ordinateur, quels sont les phénomènes que l'on rencontre vraiment ? (...) Les quelques corpus qui ont été réalisés et analysés (par exemple, [MOREL 85]) portent souvent sur des tâches restreintes pour lesquelles ces capacités sont inutiles ; en outre, ils ont été simulés (un compère, caché par plus ou moins de machinerie, jouant le rôle de l'ordinateur) et les conclusions qu'on peut en tirer ne sont pas d'une validité certaine... » [SABAH 89].*

A partir de ces considérations, nous avons cherché à mettre en œuvre un système de DHM, dans lequel les réactions et le comportement de la machine ne soient pas simulés, mais véritablement effectifs. Nous nous sommes alors intéressé au réseau Internet, comme terrain d'expérimentation pour nos recherches, afin d'obtenir rapidement et efficacement des corpus de dialogues entre des utilisateurs et une machine.

1.2 Le réseau Internet

L'étymologie du terme Internet révèle le préfixe *Inter*, qui indique un lien entre deux éléments, et le mot *Net* qui est la traduction anglaise pour «réseau». Internet est alors un lien entre deux ou plusieurs réseaux informatiques, « un réseau de réseaux ». Au sens propre, il s'agit du plus grand réseau informatique de la planète. Il regroupe une multitude de réseaux régionaux, gouvernementaux et commerciaux.

Internet est un réseau physique de télécommunications, utilisant le protocole de communication TCP/IP¹, sur lequel on retrouve de nombreuses ressources, que nous passons brièvement en revue, ci-dessous.

Le **Courrier électronique** ou **E-mail**, est sans doute l'un des services le plus utilisé par les usagers² d'Internet. Il permet l'envoi de messages sous format électronique, d'une personne, à un ou plusieurs destinataires, sur un réseau informatique. **Telnet** permet à un utilisateur d'entrer en communication avec un ordinateur distant, sur le réseau Internet. **FTP** (File Transfer Protocol) est un protocole utilisé lors d'un transfert de fichiers entre deux ordinateurs sur un réseau TCP/IP. **Archie** indexe le contenu des serveurs FTP anonymes. **Gopher** est une ressource qui permet de naviguer au sein d'Internet à l'aide de menus arborescents. Plusieurs milliers de serveurs Gophers sont utilisables sur Internet aujourd'hui. **IRC** (Internet Relay Chat) permet à des utilisateurs de s'écrire en temps réel. Ce type de communication ne transmet pas la voix mais des phrases rédigées au clavier. IRC est un système multi-utilisateurs où les gens se réunissent dans des endroits virtuels appelés *canaux* (channels en anglais) pour discuter d'un sujet donné. Les nouvelles **Usenet** (on parle aussi de *News*) sont un ensemble de quelque 10 000 groupes de discussions qui traitent de sujets variés. Tous les internautes peuvent lire et envoyer des articles (messages électroniques) dans les News. **Wais** est l'acronyme pour «Wide Area Information Server». Il s'agit d'un moteur de recherche et

¹ TCP/IP est le sigle de "Transfer Control Protocol/Internet Protocol".

d'indexation qui fonctionne sur un réseau TCP/IP. Les bases de données interrogeables grâce à cet outil sont généralement institutionnelles. Les banques Wais peuvent être consultées par un client du même nom, ou bien par le biais du **WWW** ou de Gopher.

➤ **Le World Wide Web**

Le World Wide Web, la « toile d'araignée mondiale », nommé plus brièvement le WWW, le W3 ou le Web, a été défini à partir de 1989 au CERN (Centre Européen de Recherche Nucléaire) de Genève, par une équipe dirigée par Tim Berners-Lee. L'objectif initial consistait à faciliter les échanges, par le réseau Internet, de documents (textes et images) des chercheurs de physique des particules, entre les différents centres de recherche dans le monde. Un nouveau type de logiciel a ensuite été créé pour automatiser ces échanges sur le Web. Appelés navigateurs (*butineur* au Canada, *browser* en anglais), ces logiciels, très faciles à utiliser et quasiment gratuits, ont permis de populariser l'accès aux documents du Web, et aux divers services offerts sur Internet³.

Sur Internet, les bases de données sont reliées entre elles par l'intermédiaire de pointeurs, ce qui permet l'accès à une grande variété de données : textes, son, vidéo, etc. « *Tim Berners-Lee et Robert Cailliau avaient décidé d'abandonner les arborescences hiérarchiques chères aux structures de répertoires des ordinateurs et avaient décidé de s'appuyer sur la conception d'hypertexte qu'avait développé Ted Nelson dès les années 1960.* » [GUEDON 96].

La Figure 1, ci-après, est la reproduction d'une gravure du capitaine Agostina Ramelli (Paris, 1588) intitulée : « La nouvelle et ingénieuse machine ». Elle devait faciliter les études, en permettant à un chercheur de consulter des informations ponctuelles, simplement en tournant une roue, sur laquelle un assistant fixait plusieurs livres.

² Dans ce document, nous employons les mots « utilisateur » et « usager » comme des synonymes.

³ On peut, en effet, rejoindre la quasi-totalité des ressources Internet - comme Gopher, Usenet, Wais, Ftp, Telnet - par le biais du W3. C'est pourquoi beaucoup de personnes pensent faire référence à la même chose en parlant du World Wide Web et d'Internet, alors qu'en fait, le Web n'est qu'une partie d'Internet.



Figure 1 : Le premier "browser" connu (Paris, 1588)

La notion d'hypertexte a commencé à être étudiée à partir de 1945, quand Vannevar Bush, dans un célèbre article de *Atlantic Monthly*, esquissa ce que pourrait être le système Memex. Pourtant, l'*Encyclopédie* de Diderot incorporait déjà des dimensions que l'on peut qualifier, anachroniquement, d'hypertextuelles.

Avec l'arrivée de l'informatique, l'archivage électronique de documents a permis de stocker de grandes quantités d'informations. La navigation hypertextuelle permet de parcourir ces données, de manière non linéaire. Mais, en contrepartie, des problèmes liés à la navigation dans ce type de documents sont apparus. [BEAUFILS 96], par exemple, a étudié la navigation en environnement hypertextuel et a montré que l'on retrouve principalement quatre sortes de difficultés navigationnelles :

- Les problèmes liés à l'évaluation d'un objectif de consultation ;
- Les problèmes liés à la maîtrise d'un environnement particulier ;
- Les problèmes généraux d'orientation, de localisation et d'accès aux informations ;
- Les problèmes liés aux traitements des informations sélectionnées.

Ces préoccupations sont, entre autres, à l'origine du projet Orion, dans lequel s'inscrit notre thèse.

1.3 Motivation de nos travaux : Le Projet Orion

Bien avant les débuts d'Internet, les chercheurs étaient déjà confrontés aux problèmes de recherche d'informations dans de grandes masses de données. On connaît désormais d'autres problèmes liés à la navigation dans des documents à structure hypertextuelle et hypermédia. 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 KEHOE 96].

Hormis la vitesse d'accès (a), qui est du ressort des performances du matériel, les autres problèmes sont, à notre avis, liés à l'interaction Homme-Machine.

Les recherches du projet Orion sont motivées 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 (Cf. en annexe 1 quelques outils de navigation cartographique pour le Web). 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].

Le projet [ORION 99] 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 WWW. Deux pôles scientifiques travaillent ensemble sur ce projet : 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'utilisateur (cf. annexe 11 à propos du projet Orion).

1.4 Méthodologie

En 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.

Dans cette thèse, nous avons étudié l'interaction en LN pour la recherche d'informations archivées sur le Web. Au cours d'un processus incrémental (cf. Figure 2), nous avons conçu et réalisé différentes versions d'un système de DHM, directement utilisables (falsifiables⁴, démontrables) par les utilisateurs du Web.

Nous avons tout d'abord défini le domaine de l'application cible : la recherche documentaire. Nous avons observé les usages des systèmes existants, avec et sans ressources informatiques. À partir de ces observations et d'une enquête réalisée dans une médiathèque, nous avons pu mettre en évidence les freins et les motivations de ces usagers à l'utilisation d'un système informatisé d'aide à la recherche documentaire. Sur cette base, nous avons spécifié un modèle de l'utilisateur.

En nous appuyant sur les travaux publiés dans le domaine, nous avons élaboré un premier système (dit n° 0). Les usagers du Web ont utilisé ce système, hors laboratoire, l'on critiqué et commenté. Une analyse du corpus ainsi recueilli a permis de déterminer ce que les utilisateurs disaient (quoi, comment, pourquoi) et ce que la machine

⁴ Théorie ou hypothèse falsifiable : pour laquelle on peut imaginer une expérience, un fait, qui en démontre la fausseté.

comprendait et répondait. Nous avons ainsi spécifié le modèle de langage relatif à notre application.

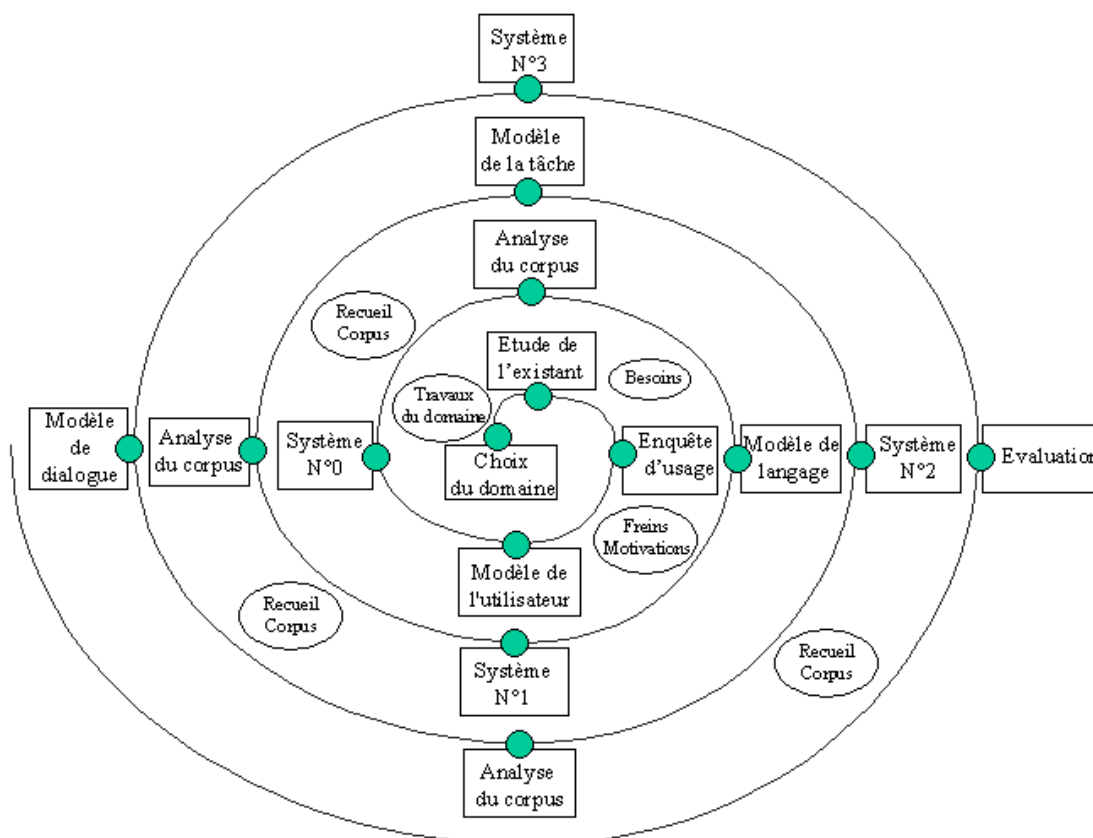


Figure 2 : Méthodologie mise en œuvre pour la conception de notre système de DHM

Le premier système (Halpin-Recueil), une sorte de jeu de devinette interactif en LN, a principalement servi de base linguistique pour l'élaboration du deuxième. La version suivante (n° 1), était un système opérationnel d'aide à la recherche documentaire. Le corpus recueilli a contribué à l'amélioration du modèle de la tâche. De même, avec la version suivante, le modèle de dialogue a été affiné, pour que le système interprète de manière plus robuste les énoncés des utilisateurs.

Le système Halpin-Documentaire est donc un système de DHM en langue naturelle, avec lequel les usagers du Web peuvent interagir de manière multimodale pour effectuer une recherche documentaire dans une base de données contenant plus de 83000 références. Ce système les aide, compte tenu de leurs besoins et de leur profil d'utilisateur. Nous verrons que son utilisation permet, dans certaines conditions, d'obtenir de meilleurs résultats qu'avec une interaction classique.

Plan du document

Après une présentation de l'état de l'art des domaines du DHM et de la recherche d'informations sur le Web (**chapitre 2**), nous exposons au **chapitre 3**, les besoins des utilisateurs en situation de recherche d'information.

Les problèmes d'amorçage d'un réel système de DHM sont abordés au **chapitre 4**. Nous décrivons alors la notion d'hyperdialogue et un système générique de DHM sur le Web est présenté.

Puis, le premier système que nous avons mis au point pour recueillir un corpus de DHM sur le Web (Halpin-Recueil) est détaillé et quelques résultats significatifs sont commentés au **chapitre 5**.

Dans le **chapitre 6**, nous présentons notre réalisation logicielle complète d'un DHM, avec entrées et sorties vocales via le Web (le système Halpin-Documentaire), ainsi que les différentes composantes nécessaires pour sa mise en œuvre sur Internet.

Le **chapitre 7** est consacré aux résultats, à l'évaluation, et aux limites du système Halpin-Documentaire.

Enfin, le **dernier chapitre** dresse un bilan de ces travaux et présente des perspectives de recherche pour le domaine étudié.

2 Situation du problème : Dialogue et Recherche d'Information

Dans les paragraphes qui suivent, nous présentons succinctement l'état de l'art dans les domaines du dialogue Homme-Machine et dans celui de la recherche d'information sur Internet.

2.1 Les systèmes de DHM

Certes, les interfaces par navigation au sein de menus sont faciles à utiliser pour des tâches relativement simples. Mais elles ne sont pas assez flexibles pour résoudre des tâches complexes, notamment lors de recherches d'information dans des bases de données : *« Pour toutes les applications informatiques simples, où l'ensemble des ordres à donner à la machine est facile à énumérer et à mémoriser, les interfaces à boutons ou à menus déroulants sont certainement supérieures au dialogue en langue naturelle. En revanche, les difficultés insurmontables des usagers dans l'utilisation des formules logiques d'interrogation de bases de données documentaires incitent à essayer d'autres méthodes. »* [NICOLLE et LUZZATI 99].

Les langages artificiels de commande, eux, sont efficaces, mais trop inabordables pour les non informaticiens. [BECKER et al. 92] illustrent cela en comparant la requête en langue naturelle anglaise *« Which sewage treatment plants have a chemical purifying stage ? »*⁵ avec son équivalent formulé en requête SQL⁶ (cf. Figure 3).

⁵ Quelles installations de traitement d'eaux d'égout ont une étape chimique d'épuration ?

```

SELECT DISTINCT
x2.OBJ_ART,x2.OBJ_NR,X1.OBJEKTNAME,X3.BEZEICHNUNG
FROM
    BECKER.WWAO x1,BECKER.VABWASSREINANL X2 ,BECKER.AWO4 X3,
    BECKER."ABWREIN-BAUWDAT " X4, BECKER.ABWASSERBAUWDATEN X5
WHERE X3.BEZEICHNUNG
IN (SELECT X6.BEZEICHNUNG
    FROM BECKER.AWO4 X6
    WHERE x6.SCHL_ID
    IN (600,120,100>
    AND X6.BEZEICHNUNG IS NOT NULL)
    AND X3.BEZEICHNUNG
    IN (SELECT X7.BEZEICHNUNG
    FROM BECKER.AW04 X7
    WHERE x7.SCHL ID = 600
    AND X7.BEZEICHNUNG IS NOT NULL)
    AND X5.ABWASSBAUWDAT ID = X4.ABWASSBAUWDAT_ID
    AND X5. "BEREICH/STUFE" = X3.SCHL_ID
    AND X4.ABWREINANL_ID = X2.ABWREINANL_ID
    AND X2.OBJ_NR = X1.OBJ_NR
    AND X2.OBJ_ART = X1.OBJ_ART

```

Figure 3 : Code SQL équivalent à la requête en LN
« Which sewage treatment plants have a chemical purifying stage ? »

De ce fait, des recherches sont effectuées dans différents domaines d'application (militaire notamment, cf. annexe 3), pour intégrer une composante dialogique en langue naturelle à certains systèmes, afin de faciliter leur utilisation.

Depuis les années 60, de nombreux systèmes de DHM ont été créés. Dans la cadre de cette thèse, nous nous sommes uniquement intéressé aux systèmes de DHM finalisés, c'est-à-dire, orientés par la réalisation d'une tâche ou la résolution d'un problème, et fondés sur l'utilisation d'une langue naturelle. Les systèmes que nous allons présenter ci-dessous ne constituent pas une liste exhaustive, mais donnent des éléments pour apprécier les spécificités de chacun d'entre eux.

Il existe plusieurs manières de répertorier les systèmes de DHM.

Colineau, dans sa classification, identifie certains des paramètres qui définissent le type d'interaction traité. « *Les principaux paramètres retenus sont ceux relatifs : à l'application, aux utilisateurs, aux connaissances que possède le système, au type de dialogue géré et au degré de naturel de la langue utilisé.* » [COLINEAU 97].

⁶ Structured Query Language

Pierrel et Romary distinguent trois grandes classes d'applications, en fonction du type d'interaction entre l'utilisateur et le système [PIERREL et ROMARY 97].

- *Le transfert ou la saisie d'information* : le dialogue en langue naturelle remplace l'utilisation d'un langage formel avec une base de données ou de connaissances.
- *La commande (ou coordination) d'action* : le système de dialogue doit jouer un rôle d'intermédiaire, pour déterminer les actions à faire et les objets référencés, afin de commander un système automatisé.
- *L'acquisition de connaissances* : le dialogue est utilisé pour la co-construction d'un modèle de référence, commun entre les interlocuteurs.

Luzzati, quant à lui, considère trois types de tâche pour la classification des systèmes :

- *Les tâches opératives (au sens de Falzon [FALZON 89])* : Une consultation de base de données, comme les renseignements d'horaires SNCF, a pour caractéristique d'être une tâche opérative, dans un domaine restreint. Les utilisateurs ont une bonne connaissance du type d'information qu'ils recherchent (horaires, numéros de train, tarifs, par exemple). La recherche peut être effectuée par un tiers : les résultats ne sont pas dépendants de l'utilisateur. Symboliquement, le système est un *esclave* : il se contente d'extraire des informations d'une base de données, et de les présenter à l'utilisateur.
- *Les tâches interprétatives* : Une recherche dans les pages jaunes est une tâche interprétative. Le domaine couvert est beaucoup plus large, et nécessite souvent, une représentation du monde, de la part de la machine. Cette dernière doit être capable d'interpréter une demande, et de maîtriser le contenu intentionnel de l'utilisateur. Les résultats sont « personnalisés », et l'utilisateur concerné peut difficilement être remplacé par un tiers, lors du dialogue. Symboliquement, le système est un *serviteur* : il doit analyser, interpréter, voire négocier avec l'utilisateur pour pouvoir le servir au mieux.

- *Les tâches réflexives* : Dans une tâche réflexive, le système mène le dialogue, et guide la réflexion. Il devient l'acteur principal. Les sujets abordés sont vastes, et peuvent évoluer au fil de la conversation. Les phénomènes d'incompréhension sont repérés et gérés par la machine. Symboliquement, le système est le *maître* : il prend l'initiative du dialogue, questionne l'utilisateur pour parvenir à ses fins.

Les possibilités de classer les systèmes sont multiples. Cependant, à notre connaissance, aucune ne prend en compte le support technique sur lequel s'effectue le dialogue. Pour notre part, nous avons choisi de classer les systèmes de DHM existants dans un espace à trois dimensions : UTS (pour Utilisateur, Tâche, Support).

* * *

1. L'axe de l'utilisateur

Il s'agit de repérer sur cet axe le degré d'adaptation du système à l'utilisateur. On considère les systèmes non-adaptables, adaptables, et adaptatifs [PIERREL et ROMARY 97].

1) Systèmes non-adaptables (dépourvus de toute fonctionnalité d'adaptation)

ELIZA est un système de dialogue, en langue naturelle anglaise, développé par Joseph Weizenbaum au MIT entre 1964 et 1966. Il simule un psychologue rogière dont le rôle est de faire parler le patient en lui retournant ses propres paroles. Le système ne s'adapte pas à son interlocuteur. Il repère des mots-clés dans les énoncés de l'utilisateur afin d'élaborer une réponse prédéfinie. Par exemple, si l'utilisateur utilise un mot comme *père, mère, ou frère*, la machine donne une réponse du genre « Dites m'en plus sur votre famille ». Le but du système n'est pas de répondre à des questions, mais au contraire, d'en poser à l'utilisateur. Ainsi, le dialogue n'est jamais bloqué, progresse toujours, et donne l'illusion d'une véritable conversation entre un patient et son psychiatre.

ERMA, développé en 1977 par Clippinger est un système similaire à Éliza. « *Ce système simule des erreurs de performances lors d'un dialogue, en anglais, avec un psychiatre.* » (cf. [PONTON 96]⁷). Il ne s'adapte pas non plus à son interlocuteur.

SHRDLU, de Terry Winograd, date de 1972. Le but du système est de répondre à des questions de l'utilisateur sur un monde clos, composé d'une table, sur laquelle sont posés des polyèdres de tailles et de couleurs diverses. Afin de dialoguer, il utilise un module de syntaxe, un autre de sémantique, et un dernier pour la résolution de problèmes. « *Comme ELIZA, ce système appartient à la famille des générateurs « à trous ».* Il détermine le centre d'intérêt de la question (*focus*) qui lui permettra par la suite de déterminer la forme prédéfinie la mieux adaptée. » [PONTON 96].

Pour Pierrel et Romary, c'est le premier système à aborder des questions réellement intéressantes pour le dialogue : « *Il pouvait mener un dialogue plus complet qui ne se limitait pas à un simple échange, expliquer ses actions et poser des questions lorsque les indications fournies étaient insuffisantes.* » [PIERREL et ROMARY 97].

2) Systèmes adaptables (paramétrisation d'un modèle statique de l'utilisateur)

Le système **BLAH**, de Weiner, écrit en 1980, n'est pas un véritable système de DHM, car les entrées sont des données brutes (fichiers et déclarations d'impôts). Cependant, nous le citons comme exemple de système adaptable à l'utilisateur. En effet, il utilise un module statique de l'utilisateur afin de ne pas expliquer à ce dernier ce qu'il sait déjà. (cf. [PONTON 96]).

ROMPER (Responding to Object-related Misconceptions using PERspective) est un système de DHM, en anglais, développé par Mc Coy en 1985. Il réagit en fonction des contradictions apparaissant entre ses connaissances et celles de l'utilisateur. Les réponses du système sont fonction du profil de l'utilisateur. Mais les caractéristiques de

⁷ Dans sa thèse sur la « *Génération de textes en langue naturelle* », Ponton a référencé 93 systèmes, dont quelques uns ont pour cadre le DHM.

l'utilisateur sont données en début de session, et non pas calculées au cours du dialogue [MC COY 85].

3) Systèmes adaptatifs (modèle dynamique de l'utilisateur en fonction de la situation et du type d'utilisateur)

MICS, de Carbonnell, 1981, est un système de DHM en langue anglaise. « Il est capable de mener un dialogue avec quelqu'un qui lui est complètement inconnu au départ. Il possède pour cela des connaissances sur les buts de certaines classes de personnes (i.e. étudiants, ...). Il utilise des règles de conversation particulièrement étudiées pour les applications à la création et à la compréhension des conversations. » (cf. [PONTON 96]). La prise en compte de l'utilisateur se fait donc, au fil du dialogue, grâce à un modèle dynamique.

GOSSIP a été écrit en 1988 par L. Jordanskaja et D. Carcagno. Gossip signifie en anglais bavarder, raconter des ragots, faire des commérages. Ce système bavarde donc avec l'utilisateur, et fournit des résumés des dialogues antérieurs. « *Pour cela, GOSSIP enrichit au fil du dialogue une base de données relative à l'utilisation et à l'utilisateur. Cette base contient entre autres des informations du type : nom de l'utilisateur, date, durée d'utilisation, type de manipulation ... D'autre part, GOSSIP est capable d'inférer de nouvelles informations à partir d'informations existantes, ceci permet de produire un discours adapté à la demande.* » (cf. [PONTON 96]).

KAMP signifie Knowledge And Modalities Planner [APPELT 81]. Ce système génère, dans le cadre d'un dialogue avec l'utilisateur, des répliques qu'un expert aurait pu produire pour aider à l'assemblage ou à la réparation d'une pièce d'équipement. « *C'est l'un des premiers systèmes à utiliser réellement une modélisation des croyances.* » [SABAH 89].

COMPEROBOT [VIVIER et NICOLLE 97] et **CEDRE** [ANDRÈS 95] sont des systèmes de dialogue, en français, dans lesquels les utilisateurs sont principalement des enfants de 8 à 11 ans. « *Afin d'établir un dialogue verbal homme-machine, avec une composante*

dialogale, il faut que la machine se fasse une image de son interlocuteur pour reconnaître ses intentions et ses stratégies. C'est ce qui fait la différence entre la communication homme-machine de type stimulus-réponse, ayant seulement un aspect dialogique, et une dialogue homme-machine plus complet, avec une composante dialogale [VERNANT 94].».

* * *

2. L'axe de la tâche

Dans sa typologie des tâches, Luzzati propose 6 classes de systèmes, en fonction du type de consultation à la base de données (BD) liée au système [LUZZATI 89]. Sur notre axe de la tâche, nous reprenons cette typologie, et ajoutons la classe 0, pour les systèmes n'accédant pas à une BD.

0) Tâche sans consultation de BD

POLITICS est une simulation de discussions politiques, en anglais. Ce système, développé par Carbonnell, en 1977, interprète des événements géopolitiques et y réagit selon des idéologies variées. Il comprend des énoncés comme « *Les russes ont massé des troupes le long de la frontière tchèque.* », et peut s'exprimer selon plusieurs points de vue : conservateur, communiste, etc. Son but principal n'est pas de fournir à l'utilisateur des données puisées dans une BD, mais de soutenir un discours argumenté en langue naturelle. (cf. [PONTON 96]).

Dans certains systèmes de dialogue orientés par la tâche, le langage permet de commander des instruments ou des machines. [MATROUF et al. 87] expliquent, à propos de leur système permettant de traiter un dialogue oral entre des élèves contrôleurs aériens et un simulateur de trafic aérien : « *Le système a pour tâche, d'une part de poursuivre un dialogue avec l'élève et de communiquer les désirs de ce dernier au simulateur sous forme de suites d'instructions du langage formel de commandes, et d'autre part, de détecter et corriger les erreurs dues aux différentes étapes de la compréhension du message.* ».

Les didacticiels comme KDS, CHESSIA, GEODE, ou les éditeurs graphiques multimodaux comme LIMSI-DRAW ou ICP-DRAW sont aussi des systèmes de DHM sans accès à une BD.

LIMSI-DRAW est un éditeur graphique multimodal développé dans une optique exploratoire. Il permet de créer et de manipuler des formes géométriques simples (rectangles, triangles, cercles) à travers un ensemble de commandes multimodales. Il intègre trois périphériques en entrée : un système de reconnaissance vocale, un écran tactile et une souris. La multimodalité supportée par le système est du type synergique. L'utilisateur dispose de plusieurs possibilités pour exprimer ses commandes : souris ; écran tactile ; souris et écran tactile ; parole et souris ; parole et écran tactile ; parole, souris, et écran tactile [BELLIK et TEIL 92]. L'utilisateur peut prononcer la phrase « *mets ça ici* », par exemple, tout en désignant un objet puis sa nouvelle position., ou encore « *efface* », en désignant un objet (par la souris ou l'écran tactile), etc.

L'éditeur graphique **ICP-DRAW** [WRETO et CAELEN 89] est une application de dessin à manipulation directe, dans lequel la communication homme-machine est multimodale. L'utilisateur dispose d'une palette d'outils graphiques et de menus de fonction. Il peut activer ces fonctions au clavier, à la souris, ou par la parole. On peut dire par exemple « dessine un cercle noir », « dessine un carré vert », etc.

KDS signifie Knowledge Delivery System. Il a été écrit en 1981, par William C. Mann et James A. Moore. Ce système, qui utilise la langue naturelle anglaise, explique à son interlocuteur « quoi faire en cas d'incendie » [MANN et MOORE 81].

CHESSIA est un système français d'EIAO (Enseignement Intelligemment Assisté par Ordinateur), qui enseigne à un utilisateur les stratégies du jeu d'échecs, grâce à un dialogue structuré et à des icônes en forme de visage, symbolisant les émotions (satisfaction, étonnement, déception, ...) [GRIMALDI 93].

GEODE [BOUFAIDIA 93] est un système destiné à la génération des interactions avec l'apprenant. Le dialogue avec le système offre des activités spécifiques à l'enseignement (aide et explication, entre autres). L'interface est indépendante de la gestion du dialogue et permet ainsi de représenter n'importe quelle matière à enseigner.

1) Tâche avec consultation de BD explicite et limitée

Tâche qui peut être effectuée par un tiers (renseignements horaires de transports, renseignements téléphoniques, etc.).

VOYAGER peut engager des conversations concernant une région géographique limitée, autour de Cambridge (USA). L'utilisateur peut obtenir des informations sur les banques, les bibliothèques et autres services disponibles, grâce à un système multilingue, anglais, italien, japonais [GLASS et al. 95].

Avec **WAXHOLM** [BLOMBERG et al. 93], l'utilisateur peut consulter des horaires de ferries pour des archipels de Stockholm. Le dialogue permet également d'obtenir des informations à propos des ports, des hôtels, des campings et des restaurants répertoriés.

LOADSTAR [HUANG et al. 99] est un système de DHM pour une recherche d'information sur les voyages, en Chine. Il peut donner des informations touristiques en mandarin, et des itinéraires en fonction des préférences des usagers, du trafic constaté, de la durée, du coût, etc. Dans ce système, l'initiative peut être prise par la machine ou par l'humain.

2) Tâche avec consultation de BD explicite et limitée + action

Tâche qui peut également être effectuée par un tiers (réservation train/avion, billetterie automatique, achats par correspondance, etc.)

GUS (Genial Understanding System) a été mis au point au centre de recherche Xerox de Palo Alto [BOBROW et al 77]. Son but est de gérer des DHM réalistes dans le cadre de la réservation de billets d'avion. Avec GUS, l'initiative est partagée entre le système et l'utilisateur. Il interprète certaines réponses indirectes de l'utilisateur et prend en

considération certaines anaphores et ellipses. Le dialogue permet de consulter la BD, et d'effectuer des actions comme la réservation, la modification ou l'annulation de vols.

DIANA [ANGLESSY 89] est un système d'interrogation en LN semblable à GUS, qui fonctionne par détection de mots-clefs. L'utilisateur dialogue en français, pour obtenir des indications sur les horaires et les tarifs de trains.

CRRS (Conference Room Reservation System) est un système québécois de réservation de salles de conférence, par téléphone [PATERAS et al. 99]. Il s'inspire des tâches d'organisation de réunions, comme dans les systèmes **PAL** [BULLWINKLE 77] et **NUDGE** [GOLDSTEIN et ROBERTS 79]. Il permet de réserver ou d'annuler des salles, en dialoguant en langue naturelle anglaise, et met à jour la base de données, au fur et à mesure des interventions. Il demande un complément d'information si nécessaire. Il offre également des alternatives lorsque les demandes initiales ne peuvent être satisfaites.

Des logiciels semblables ont été développés par AT&T. Ils se basent sur une architecture conversationnelle à initiative mixte, nommée **AMICA**. « *In fact the same architecture has been used for prototypical implementation of the ATIS applications, a conference room reservation, a restaurant and a movie information task.* » [PIERACCINI et al. 97].

PEGASUS [ZUE et al. 94] permet d'obtenir des données relatives aux voyages aériens. Ce système, consultable par téléphone, connaît plus de 220 aéroports dans le monde. Il peut aider les usagers dans différentes situations : « *It has a fairly extensive dialogue component to help cope with difficult problems such as date restrictions imposed by discount fares or aborted flight plans due to selections being sold out.* ».

3) Tâche avec consultation de BD explicite non limitée

Tâche qui fait appel à de vastes connaissances sur le monde (consultation de pages jaunes, etc.)

WEBGALAXY est un système conversationnel multi-domaine : météo nationale, pages jaunes, et accès aux informations du Web [LAU et al. 97]. La communication homme-machine se fait principalement par téléphone : « *By utilizing the telephone network for speech input and output, and Web protocols for the graphical interface, the WebGALAXY architecture advances the SLS goal of enabling universal access to online information.* »⁸. On peut lui poser des questions comme « *Quel temps fera-t-il demain à Miami ?* », « *Combien y a-t-il d'hôtels à Boston ?* », ou encore « *As-tu des informations sur la Suisse ?* » [ZUE 97]. Ce système accède donc à des bases de données non limitées (comme le Web), et permet une interaction multimodale. L'utilisateur peut, par exemple, demander par téléphone « *Je veux atteindre la page Web du MIT* », et la page principale de ce laboratoire s'affiche à l'écran. On peut alors cliquer sur les liens comme habituellement. Le système n'est cependant pas accessible au grand public : « *Due to limited resources, WebGALAXY is not available to the public at this time.* ».

4) Tâche avec consultation de BD implicite

L'utilisateur ne sait pas *a priori* ce qu'il peut trouver dans ce genre de BD (renseignements professionnels, Centre d'Information et de Documentation, etc.)

TEXT, de Kathleen R. Mc Keown, est un système de DHM qui est relié à une base de données militaire [MC KEOWN 82]. Il a pour objectif de répondre à des questions relatives aux véhicules militaires et aux armes utilisés dans la marine. Contrairement à une base de données de type SNCF, pour laquelle l'utilisateur connaît les arguments sur lesquels il peut effectuer des requêtes (tarif, horaires, correspondances), ici, l'utilisateur doit, en quelque sorte, « découvrir » la BD, grâce au dialogue : il sait que l'information existe, mais ignore souvent sous quelle forme elle est disponible.

DIAL [ROUSSANALY 88] est un système de communication orale homme-machine, servant à consulter des informations administratives, en langue naturelle. [ROUSSANALY et PIERREL 92] précisent de quel type d'application il s'agit : « *Au sein de notre équipe,*

nous avons choisi d'aborder deux familles d'applications assez éloignées et donc de recourir à deux type distincts d'interfaces orales :

- d'une part, des applications classiques où l'ordinateur intervient comme exécutant et qui mettent en œuvre un univers pragmatique restreint comme la commande de processus : système **PARTNER** [MORIN 94] et système **DIAPASON** [SOUVAY 92].
- d'autre part, des applications concernant des activités cognitives plus complexes dans un univers plus riche comme les systèmes d'aide appliqués ici au renseignement administratif : système **DIAL**. »

MIRACLE (Multimedia Information Retrieval of Concepts in a Logical Environment), **MERIT** (Multimedia Extension of Retrieval Interaction Tools) et **CORINNA** (Cooperative Retrieval Interface for Natural Language Acts) sont des systèmes qui permettent d'obtenir des informations, en découvrant, au fur et à mesure du dialogue, le contenu de la BD. Les domaines d'applications sont vastes (histoire de l'art, information sur les projets de la communauté européenne, ...). Dans **MIRACLE**, les questions posées par les utilisateurs peuvent être du genre « *search for artists concerned with abstract art in Spain* » [FISCHER et al. 94]. Quand l'information n'est pas directement accessible, le système réoriente le dialogue. Il est capable d'expliquer son raisonnement, si besoin. A la question « *Quels sont les projets ayant dans leur titre NL GENERATION ?* », **CORINNA** répond qu'il a besoin d'un terme plus général pour lancer une requête. Si on lui demande pourquoi, il explique que c'est parce que ce terme là n'est pas connu dans la BD. Avec ce type de systèmes, l'utilisateur découvre, au fil du dialogue, les actions qu'il peut effectuer et les informations auxquelles il peut accéder.

5) Tâche avec consultation de BD supposant de la part du système des facultés de contestation (réclamations)

« **BISTRO** est une simulation d'un distributeur de boissons qui dialogue en langue naturelle et qui apprend dans et par le dialogue » [GERARD et NICOLLE 98]. Lorsqu'il ne

⁸ <http://www.sls.lcs.mit.edu/sls/whatwedo/applications/webgalaxy.html>

connaît pas la recette d'une boisson, il la demande et l'apprend (en reformulant si nécessaire). Il est ensuite en mesure de faire référence à cette recette dans d'autres dialogues.

6) Tâche avec consultation de base de connaissances, c'est-à-dire d'une forme complexe de BD (documentation automatique, expertise médicale)

Une base de connaissances est un ensemble d'informations, parfois disponibles sous forme d'articles, qui répondent à une demande très précise. Elle peut être mise à jour fréquemment, avec des notions de classe et d'objets, d'attributs et de méthodes, d'agrégation et d'héritage.

[VILLASEÑOR-PINEDA 97] a amélioré le système **ICP-DRAW** – présenté plus haut – grâce, notamment, à un processus d'interprétation d'actes de dialogue. Dans son système, la machine apprend au fil du dialogue, et tente d'apporter son aide à l'utilisateur dès qu'elle reconnaît une tâche. Par exemple, l'utilisateur demande à l'ordinateur de dessiner un carré, puis un triangle. La machine demande alors « *es-tu en train de dessiner une maison ?* ». Si l'utilisateur confirme, la machine propose de compléter la tâche pour lui. Il est ensuite possible d'ajouter des éléments à certains objets pour les spécialiser. La machine enrichit alors sa base de connaissances.

* * *

3. L'axe du support

Cet axe détermine le support par lequel le système est accessible.

1) Ordinateur (sans réseau)

Dans **BISTRO** - que nous avons déjà présenté – le vrai support du système est un ordinateur (et non pas un distributeur de boisson). En effet, le système ne délivre pas réellement de boissons, mais dialogue avec les usagers comme s'il allait le faire vraiment.

Dans **AMI** (Agent de Messagerie Interactive), la tâche du système est double : d'une part mettre en relation une personne appelante et une personne appelée, et d'autre part

mettre en œuvre un serveur de messagerie. « *L'application choisie est assez répandue dans la communauté. Elle est similaire à celle du système **STANDIA** [GRAU et al. 94].* » [LEMEUNIER 99].

Bien que remplissant la fonction de « standardiste téléphonique », le support du système AMI est bien l'ordinateur. D'ailleurs, cette application, en partie développée avec le langage Common Lisp, se base sur l'architecture du système **COALA** [LEHUEN 97] (COAdaptation Langagière pour l'Apprentissage), qui est un système d'aide documentaire dans une médiathèque municipale.

2) Téléphone

Beaucoup de systèmes de DHM sont basés sur une communication téléphonique. C'est le cas dans le projet de renseignements ferroviaires **ARISE** (Automatic Railway Information Systems for Europe) [DEN OS et al. 99], ou encore avec **JUPITER** (*a telephone-only conversational interface to weather information for more than 500 cities worldwide*) [ZUE 97].

SUNDIAL (Speech UNderstanding and DIAlOg) est un système de dialogue oral par réseau téléphonique [PECKHAM 91]. Son architecture comporte trois composantes : le dialogueur, l'analyseur linguistique et le système de reconnaissance. Le module de la tâche permet de spécifier une application cible sans que le module de dialogue soit modifié.

RAITEL (Railway Telephone Information Service) est un système de DHM, développé au laboratoire LIMSI [BENNACEF et al. 96] permettant d'obtenir des informations concernant les horaires de trains (SNCF), via le téléphone. La voix est ici le seul moyen d'échange d'information entre les interlocuteurs. Les outils de reconnaissance de la parole utilisés dans ce genre d'applications sont donc déterminants pour le bon fonctionnement du système. RAITEL est capable de répondre à des sous-dialogues de demande d'explications, de reformuler une question s'il n'a pas tous les éléments nécessaires pour une tâche donnée, etc. Lorsque l'accès à la base de données dure un

certain temps, le système déclenche un métadialogue et génère des messages d'attentes (« *Veillez patienter SVP, nous essayons de répondre à votre requête.* »), de manière à informer l'utilisateur de l'état d'avancement de la tâche.

SAPLEN, un système en langue espagnole, a pour objectif de prendre les commandes des clients d'un Fast-Food [LOPEZ-COZAR et al. 99]. Il dispose d'un vocabulaire de 2000 mots. Les usagers dialoguent avec lui par téléphone, à propos des produits que propose le restaurant, mais il peut également les renseigner sur les rues, les avenues et les quartiers proches du restaurant.

3) Internet / Minitel

A notre connaissance, il n'existe pas de système de DHM accessible via le Minitel. Les pages de renseignements du 3611 autorisent l'interrogation en langue naturelle, mais il n'y a pas de dialogue possible avec la machine.

Sur Internet, seule une version d'ÉLIZA, développée en Java par Charles Hayden (Bell Labs), est accessible en ligne⁹. Le dialogue est uniquement écrit. L'utilisateur tape sa phrase, en anglais, dans une boîte de dialogue et le système répond dans une fenêtre faisant également office d'historique.

4) Borne interactive

MASK (Multimodal-Multimedia Automated Service Kiosk) est un système de DHM, qui se présente sous la forme d'un kiosque multimodal en entrée (parole et toucher), et multimédia en sortie (son, vidéo, texte et graphique) [LIFE et al. 96]. Installé dans une gare SNCF parisienne, son but est de fournir des informations sur les voyages, et de délivrer des tickets.

Le système **GEORAL** a pour objectif de mettre en œuvre une application multimodale utilisant le graphique en sortie, et la parole en entrée et en sortie. Le but de l'application est l'interrogation d'une base de données cartographiques et touristiques visualisées sous

forme graphique. **GEORAL TACTILE** est la continuation du Projet Géoral. Il se focalise notamment sur l'intégration d'un canal de communication supplémentaire : un écran tactile. Il interprète des actes de communication oraux et tactiles, et la communication homme-machine est fondée sur le concept de planification. Les premières observations montrent l'intérêt du tactile pour la compréhension globale des énoncés des utilisateurs. Cependant, son usage par des utilisateurs « grand public » n'apparaît pas comme très spontané [SIROUX et al. 95].

5) Robot mobile

Un système de DHM a été intégré dans un robot mobile, au Japon [ASOH et al. 99]. Lorsque ce robot rencontre une personne, une camera s'active, et le système compare le visage de l'interlocuteur à ceux stockés dans sa base de données. S'il la reconnaît, le robot salue la personne, et peut entamer une conversation avec elle (dans un environnement bruyant). Il est capable de répondre à des questions portant sur la localisation de personnes (« Q = *Où est Mr. Asoh ?* ; R = *Dans le bâtiment principal* »). Il peut guider (amener physiquement) une personne jusqu'à un point donné de l'immeuble. Il peut prendre l'initiative d'envoyer des E-mails à un membre du personnel, pour lui annoncer la venue d'un visiteur.

Il lui arrive parfois de perdre ses repères, et de ne plus savoir où il se trouve dans le bâtiment. Dans ce cas, il interroge des humains afin de recalculer son chemin, en fonction des informations fournies (« Q = *Où suis-je ?* ; R = *En face du bureau du Dr. Hara* »).

Enfin, équipé de microphones directionnels, il arrive à détecter plusieurs sources simultanées de sons et de paroles. S'il est déjà en conversation avec un interlocuteur et qu'un autre tente d'entrer en communication avec lui, il effectue un mouvement directionnel de la tourelle (représentant sa tête) vers cette personne, et dit « *Un instant je vous prie* », puis retourne à la discussion en cours. Cette dernière terminée, il continue le dialogue laissé en attente.

⁹ <http://chayden.net/chayden/eliza/Eliza.shtml>

6) Système embarqué

MINOS [SASAJIMA et al. 99] est un prototype de système de DHM embarqué dans un véhicule. Il répond aux questions de localisation (parking, magasin, etc.) en japonais. Il donne aussi des estimations de durée de déplacement entre deux points, en fonction des indications de l'utilisateur (halte par exemple) ou constatées (vitesse de la voiture et positionnement par GPS).

Le système français **CARMINAT** [RENAULT 95] fait également partie des systèmes de DHM embarqués. Il propose au conducteur un itinéraire personnalisé (le plus rapide ou le plus court). Il tient compte des éventuels problèmes de circulation et suggère alors un itinéraire bis. En plus des informations concernant le trafic routier, le système dispose de données à propos des services disponibles tels que les parkings ou les hôtels. Une carte routière et un curseur apparaissent sur l'écran de bord du véhicule. L'utilisateur peut alors effectuer des manipulations (zoom, etc.) pour interagir avec l'ordinateur.

* * *

Pour plus de rigueur, le caractère *interruptible* ou pas du dialogue pourrait être précisé pour chaque application : les interlocuteurs peuvent-ils s'interrompre au cours du dialogue, ou bien les tours de parole doivent-ils être respectés, comme en communication de type talkie-walkie ?

Enfin, la notion de retour d'information (*feedback*) est également importante dans la classification d'une application : le système permet-il à l'utilisateur de se rendre compte de certaines erreurs (mauvaise reconnaissance de la parole par exemple) ?

En fonction des trois axes, Utilisateur, Tâche et Support, un système pourra être identifié par un code UTS. Par exemple, un système ayant un code UTS 101, est un

système de DHM, non adaptable à l'utilisateur (U1), sans consultation de BD (T0), et utilisable avec un ordinateur (S1).

Sur les représentations graphiques présentées dans les pages suivantes, nous avons positionné quelques uns des systèmes de DHM que nous avons référencés. Ils ne sont pas tous représentés, d'une part pour des raisons de lisibilité des graphiques, et d'autre part, pour certains, par manque d'informations permettant de les classier correctement dans un repère UTS. Cette classification ne porte pas de jugement de valeur. On ne peut pas dire, par exemple, qu'un système de type U3 est meilleur qu'un autre de type U1, juste par le fait qu'il soit positionné sur ce plan là.

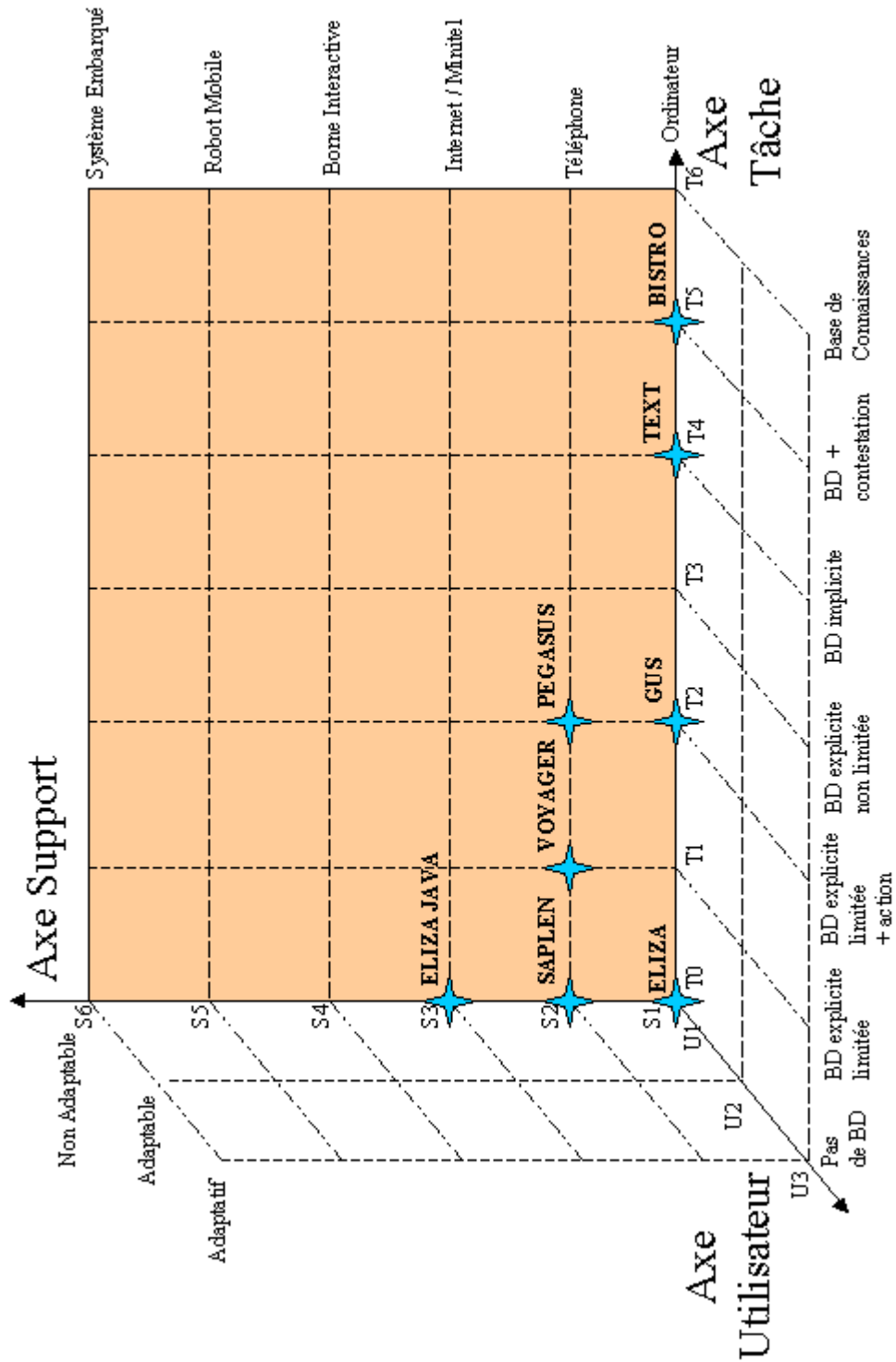


Figure 4 : Quelques systèmes de DHM de type U1

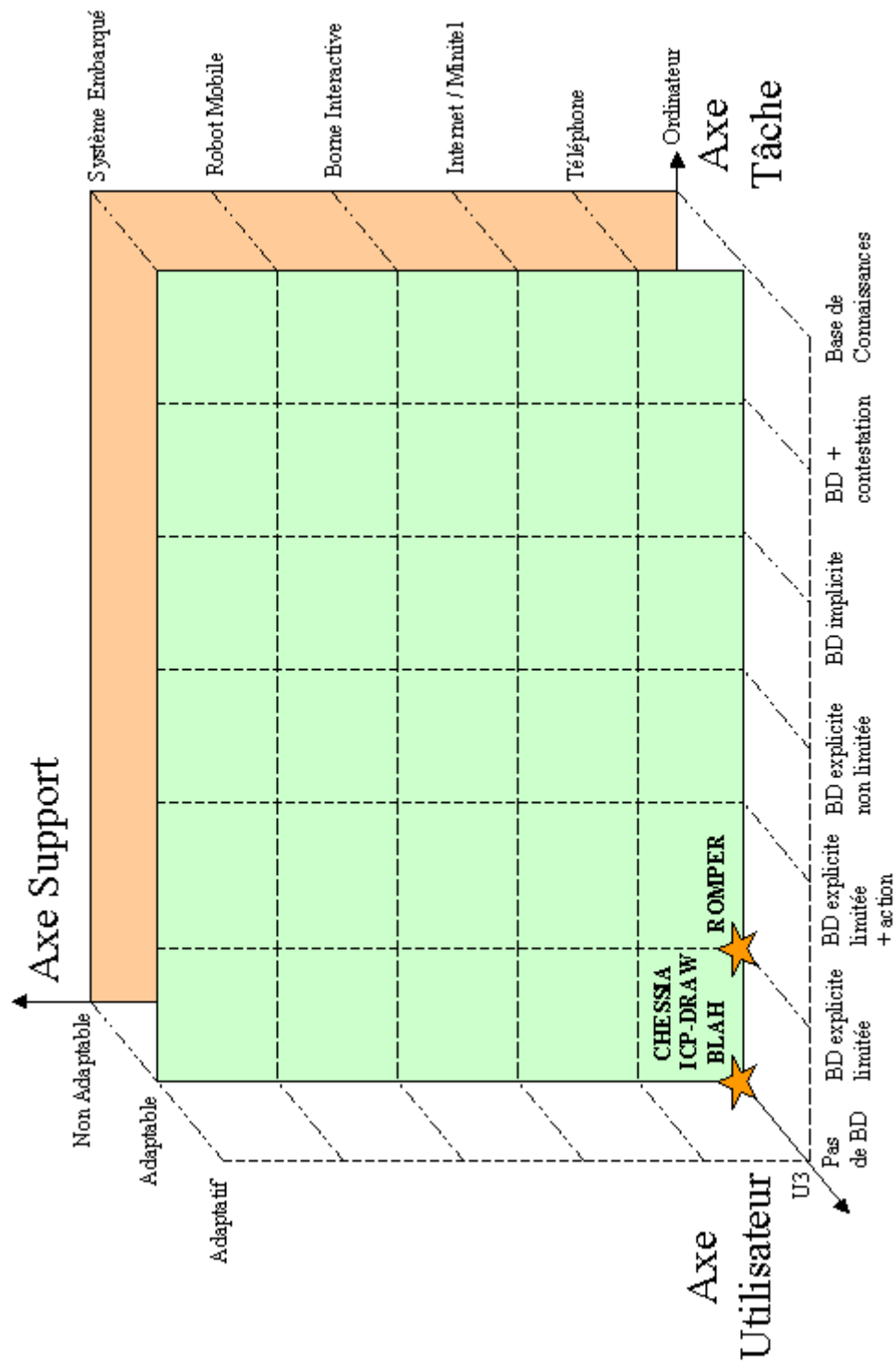


Figure 5 : Quelques systèmes de DHM de type U2

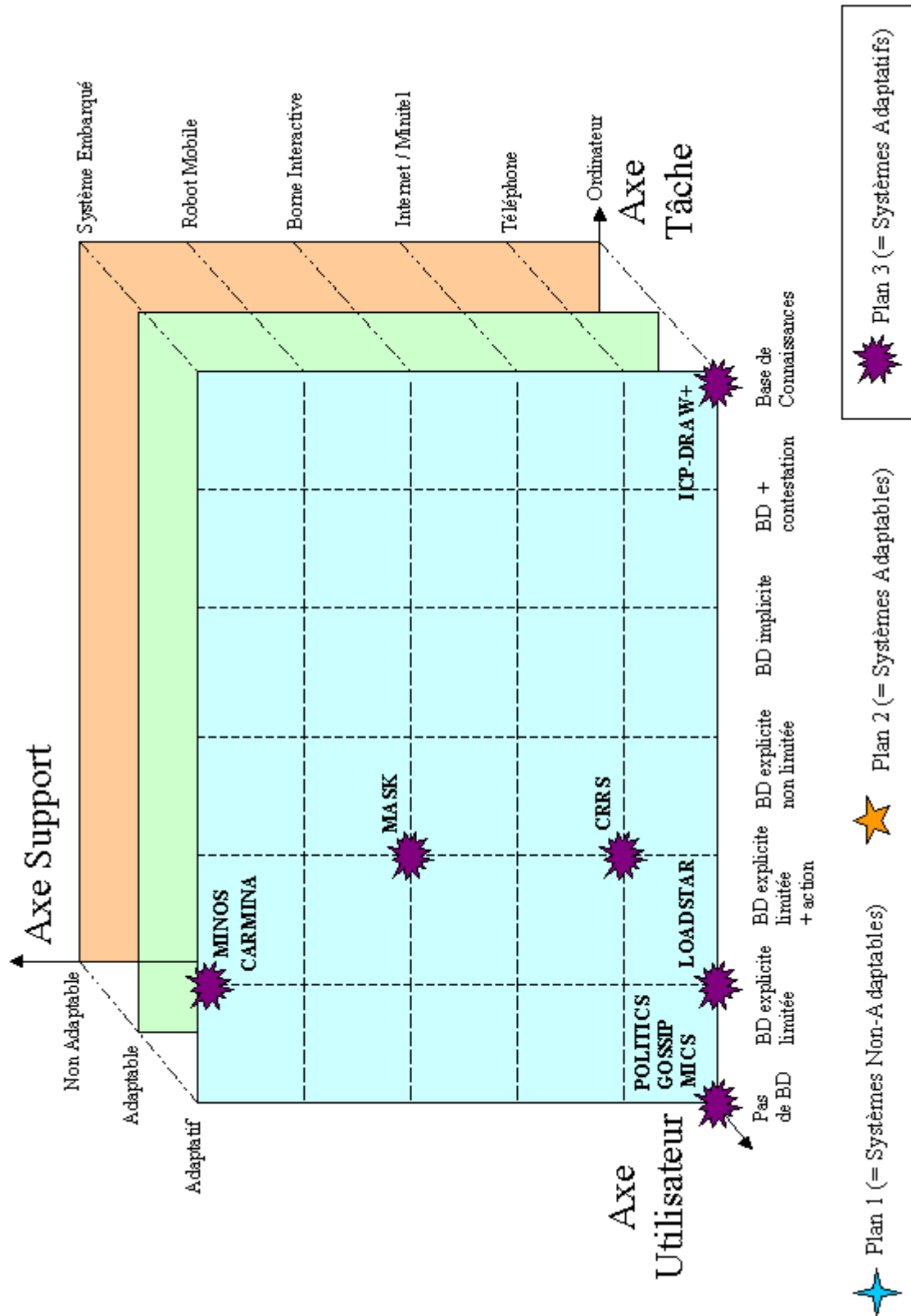


Figure 6 : Quelques systèmes de DHM de type U3

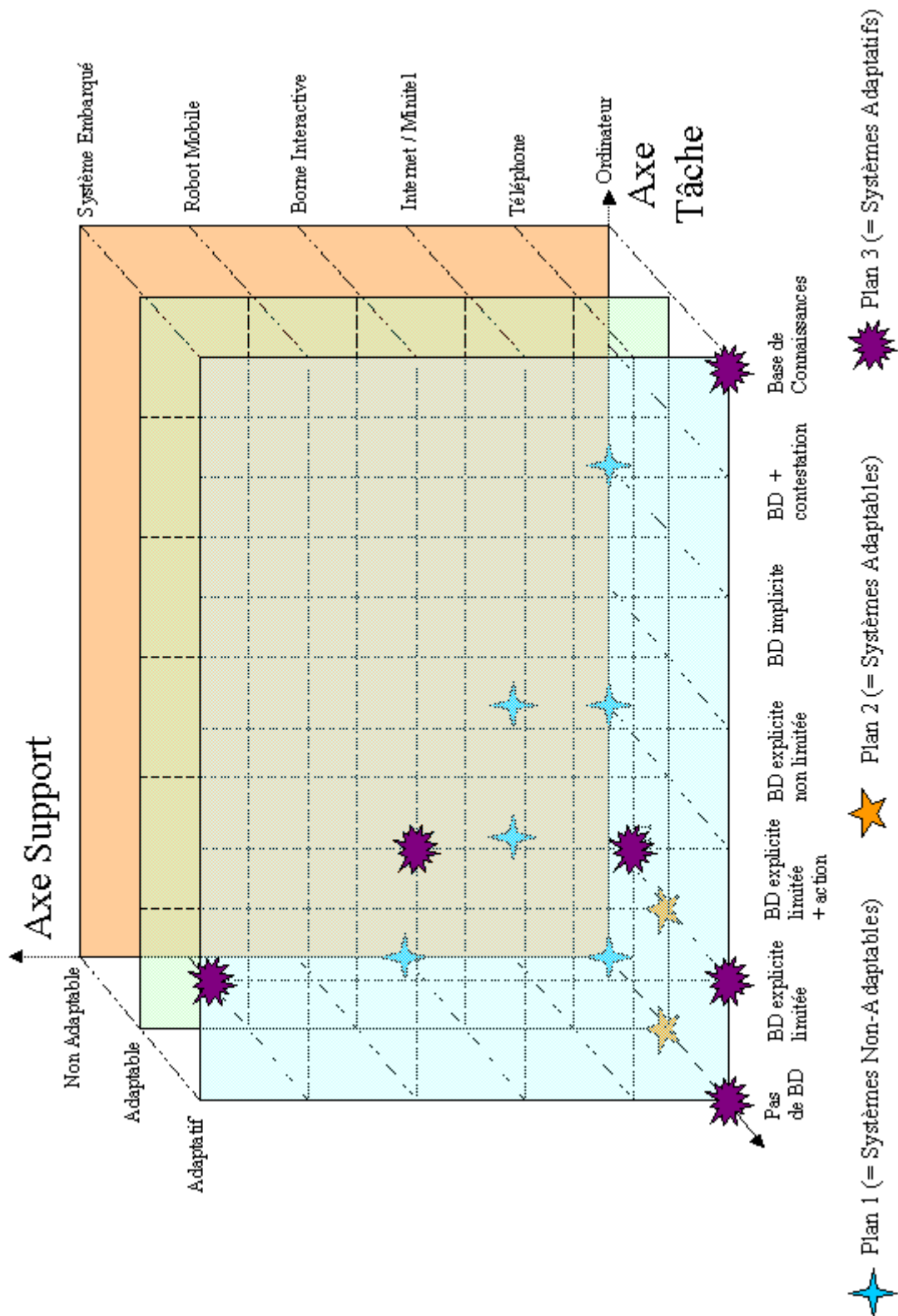


Figure 7 : Quelques systèmes de DHM positionnés sur les plans U1, U2 ou U3

Sur ces représentations graphiques, nous n'avons fait figurer que quelques systèmes existants, à titre indicatif. En schématisant, on peut dire que :

1. une très grande majorité des systèmes utilise des supports classiques, comme l'ordinateur ou le téléphone, et que très peu d'applications sont disponibles sur le réseau Internet, ou intégrées dans des robots mobiles ou des systèmes embarqués ;
2. la plupart des applications permettent d'accéder à des bases de données explicites limitées, et d'effectuer, grâce au dialogue, des actions sur ces BD ;
3. la moitié des systèmes de DHM considère l'utilisateur de manière générique, tandis que l'autre moitié des systèmes s'adapte dynamiquement à l'utilisateur, au cours du dialogue.

2.2 La recherche d'information sur Internet

La recherche d'information vise à retrouver des documents pertinents en réponse à une question posée. Traditionnellement, les documents sont indexés (manuellement ou automatiquement) avec des mots jugés importants. Au cours d'une recherche, un document sera retrouvé de manière automatique, s'il contient, dans les mots-clés qui le caractérisent, voire dans le corps du texte, les mêmes mots-clés que ceux fournis pour la requête.

Avec Internet, l'utilisateur dispose, potentiellement, du plus vaste regroupement d'informations jamais archivées. Cependant, ces informations sont disséminées sur des systèmes informatiques hétérogènes. Aucun organisme officiel n'est chargé d'indexer les documents disponibles sur le réseau Internet. De plus, dans cet environnement hypertextuel et hypermédia, un document peut être de différente nature : texte, image, son, film vidéo. « *If we compare a hypertext system to the (basic) IR system, then the major differences in terms of information access seem to be that hypertext system emphasize multimedia data, network representations, and browsing instead of queries.* » [CROFT et al. 90].

Tout le problème consiste à découvrir où se trouve l'information recherchée, et quelle est son adéquation par rapport à la question posée. Pour cela, des outils d'aide à la recherche ont été développés, notamment pour le Web. Ils suivent trois logiques de recherche distinctes : la recherche géographique, la recherche thématique, et la recherche par mots-clés (dans un index ou un méta-index).

2.2.1 Recherche géographique

L'utilisation d'un moteur de recherche géographique d'Internet permet de se focaliser, étape par étape, sur des zones géographiques de plus en plus détaillées. *« L'interface vous montre une carte ou une liste de pays et vous demande la zone désirée. La recherche se fait alors de proche en proche, sur des cartes ou des listes de plus en plus fines (pays, régions, villes), jusqu'à arriver à une compilation de services disponibles. »* [ANDRIEU 96].

Il existe plus de 80 outils de recherche géographique sur Internet. On peut démarrer la recherche de manière très large, en partant d'une carte du monde, ou bien, de manière plus précise, en commençant par le pays visé. [SAMIER et SANDOVAL 98] illustrent ce type de recherche grâce à cet exemple : *« Comment installer Alain en Écosse pour sa prochaine année universitaire à Édimbourg ? Sur City.net, sélection de la zone de l'Écosse, sélection d'Édimbourg, accès à la liste des sites de la ville sur les hébergements (un résultat en une minute). »*.

2.2.2 Recherche thématique

Basés sur le même principe de progression que les outils de recherche géographique, les moteurs de recherche thématique d'Internet permettent d'avancer, de thèmes en sous-thèmes, et d'affiner la recherche dans le domaine exploré, étape par étape. Il faut choisir une catégorie initiale parmi celles proposées par le moteur, puis approfondir la recherche pour aboutir à un résultat satisfaisant. Par exemple, pour la question *« Pierre recherche la liste des hôtels en Martinique. Sur Yahoo, il tape Martinique puis sélectionne la rubrique Hôtel. Il obtient finalement plus de 10 hôtels en 30 secondes ! »* [SAMIER et SANDOVAL 98].

2.2.3 Recherche par mots-clés

Un outil de recherche par index (*search engine* en anglais) compare les mots-clés donnés par l'utilisateur avec ceux stockés dans une base de données, qui caractérisent les pages Web référencées par ce moteur de recherche.

Les bases de données des moteurs de recherche sont mises à jour régulièrement, soit avec des informations obtenues par des robots informatiques qui scrutent le Web continuellement, soit grâce à données fournies volontairement par une personne qui veut voir son site référencé par ce moteur. À partir du titre, des mots-clés et du contenu d'une page Web, le moteur de recherche indexe cette page, de manière automatique. Ensuite, lorsqu'un utilisateur lui soumet une requête, il présente une liste ordonnée d'hyperliens, pointant vers les pages Web censées contenir l'information recherchée.

Ce type de recherche est plus efficace si l'on formule les requêtes en utilisant des opérateurs booléens (ET, OU, SAUF, etc.), afin d'ajouter ou d'exclure certains mots-clés. Une requête avec comme mots-clés *fibre ET optique* donne des résultats plus proches du domaine de la télécommunication que de ceux du tissu ou de la lunetterie.

Avec certains moteurs de recherche, d'autres éléments sont à prendre en compte pour obtenir un résultat optimum : l'ordre de saisie des mots-clés, les caractères accentués, les différences entre majuscules et minuscules, la langue, le type de document recherché.

Enfin, les méta-moteurs de recherche ne possèdent pas d'outil d'indexation propre. Ils consultent simultanément plusieurs autres moteurs et présentent les réponses en éliminant les références multiples.

* * *

Dans l'ensemble, ces outils semblent fournir des réponses convenables lorsque les questions sont précises, et que les besoins des utilisateurs sont simples et clairement définis.

Notre critique générale vis-à-vis de la plupart de ces moteurs d'aide à la recherche d'information sur Internet concerne l'inévitable nécessité de transformer la question que l'utilisateur se pose en langue naturelle, en une suite de mots-clés, qui doivent être combinés avec des opérateurs booléens.

Olivier Andrieu conclut son ouvrage consacré aux méthodes et outils de recherche sur l'Internet par ces mots : « *Il est vrai que l'on peut regretter l'absence d'un outil universel, qui serait un repère évident et automatique quel que soit l'objet de la recherche. Les notions de tri et de pertinence de l'information ne sont également pas encore réellement prises en compte dans les systèmes actuels. Mais il y a fort à parier que vont prochainement arriver des services fiables où l'on pourra poser des questions en langage naturel, de la façon la plus simple possible, et qui délivreront un maximum de 20 réponses, toutes plus pertinentes les unes que les autres.* » [ANDRIEU 96].

Enfin, les quelques sites qui proposent une interrogation en langage naturel, comme celui du CEA, ne permettent en fait qu'une formulation assez rigide¹⁰ de la requête, et surtout, ne sont pas prévus pour dialoguer avec l'utilisateur, afin de l'aider dans sa recherche.

¹⁰ Le serveur du CEA, utilisant le système Spirit précise qu'*il est inutile (sinon nuisible) de poser une requête de la manière suivante : "y a-t-il quelque chose sur le traitement des déchets nucléaires ?". On se contentera de la formulation suivante : "le traitement des déchets nucléaires"* ; [<http://www-dist.cea.fr/Spirit/aide/index.html>].

En résumé :

Au vu de ce rapide tour d'horizon des deux domaines étudiés, il nous semble important de noter, d'une part, que les systèmes de DHM utilisables sur le réseau Internet sont peu nombreux, au niveau international, voire inexistant en France, et d'autre part, que les outils actuellement disponibles pour la recherche d'information sur le Web, ne sont réellement exploitables qu'au prix d'un effort cognitif, qui consiste à traduire une question, originellement formulée en langage naturel, en une suite de mots-clés et d'opérateurs booléens, savamment combinés.

C'est pourquoi nous proposons de concevoir et de réaliser un système de DHM utilisant la langue naturelle française, et capable de s'adapter à son interlocuteur de manière dynamique, au cours du dialogue (système adaptatif de type U3).

La tâche réalisable grâce à notre système sera une consultation de base de données (système avec tâche de type T3). Plus précisément, nous souhaitons proposer l'accès à une base de données documentaire, interrogeable selon différents critères de recherche, comme le nom de l'auteur d'un document, l'année de publication, etc.

Enfin, nous souhaitons que notre système soit utilisable par un large public, et disponible sur un support autorisant plusieurs modes de communication (écrit, oral, graphique). Le réseau Internet répond à ces contraintes, et a été choisi pour ces raisons (système avec support de type S3).

* * *

Pour valider ces choix, et les caractéristiques du système que nous voulons développer, nous devons étudier les besoins et les attentes des utilisateurs dans le cadre d'une recherche d'information documentaire. Pour cela, nous avons observé et interrogé des usagers d'une médiathèque grenobloise, afin d'apporter des éléments nécessaires à la conception des modèles sur lesquels sera basé notre système.

3 Besoins des utilisateurs en situation de recherche d'information

Nous présentons dans cette partie l'étude préliminaire que nous avons menée, sur les besoins et les comportements des utilisateurs en situation de recherche d'information documentaire. Cela a pour objectif principal de repérer les connaissances initiales dont doit disposer le système de DHM. Traditionnellement, ces connaissances sont identifiées par rapport à leur caractère statique ou dynamique [LUZZATI 95].

Les connaissances statiques font référence au modèle de langage, au modèle de la tâche et à celui du dialogue ([SABAH 89], [PIERREL 87]), tandis que les connaissances dynamiques portent sur l'état de la tâche, l'historique du dialogue et le modèle de l'utilisateur.

❖ Enquête d'usage à la médiathèque de l'IMAG

Hardie a montré qu'une grande majorité des recherches des usagers sur les moteurs de recherche du Web se font à partir d'une requête à terme unique. Ce qui prédomine selon lui dans le comportement des utilisateurs d'Internet, c'est la faculté pour certains d'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 effectué une enquête auprès des usagers de la médiathèque de l'IMAG (Institut d'Informatique et de Mathématiques Appliquées de Grenoble).

Notre but consistait à étudier les utilisateurs en situation réelle de recherche d'informations ; à observer leurs stratégies, leurs méthodes d'investigation ; à leur faire critiquer les interfaces d'accès aux ressources documentaires de la médiathèque ; puis à recueillir leurs opinions concernant d'éventuelles interfaces plus anthropocentrées (LN, dialogue oral, etc.).

Nous avons élaboré un formulaire d'une vingtaine de questions (Cf. annexe 2). Puis nous avons effectué l'enquête d'usage *in situ*, durant une semaine¹¹, sur un échantillon de vingt cinq personnes.

La grande majorité des usagers de la médiathèque sont des étudiants de troisième cycle. Les diplômes qu'ils préparent sont à forte dominante scientifique : DEA Mathématiques Appliquées ; DEA Informatique (Systèmes et Communications) ; Ingénieur Ensimag ; Ingénieur INPG, etc. L'âge moyen des personnes interrogées est de 26 ans. La personne la plus jeune a 22 ans et la plus âgée a 45 ans. On dénombre 60% d'hommes. Cela ne veut pas forcément dire qu'il y a plus d'hommes que de femmes qui viennent effectuer des travaux à la médiathèque. Ces pourcentages sont sans doute égaux à la proportion d'hommes et de femmes présents sur le campus universitaire. Soixante seize pour cent des personnes interrogées sont d'origine française. Les autres nationalités rencontrées lors de l'enquête sont les suivantes : italienne, libanaise, allemande, mexicaine, algérienne et syrienne. On a observé que les usagers de la médiathèque, quelle que soit leur origine, n'ont pas de problèmes quant à l'utilisation du français tant à l'oral qu'à l'écrit. Les quelques personnes qui considèrent n'avoir qu'un niveau moyen de compréhension à l'écrit, signalent que c'est souvent lorsque le vocabulaire est trop technique, voire familier, qu'elles rencontrent des difficultés. On pourrait également considérer que seules les personnes ayant une bonne maîtrise du français viennent à la médiathèque ; les autres essayant de trouver des informations dans leur langue maternelle (sur Internet par exemple).

¹¹ Entre le 25 et le 30 novembre 1996.

Les résultats et commentaires sont ceux dont les usagers ont conscience ; il ne faut donc pas en tirer des conclusions trop hâtives. En effet, ce que les usagers «disent qu'ils font» n'est pas obligatoirement ce qu'ils font réellement¹².

❖ Résultats

Nous avons essayé de distinguer les informations obtenues en fonction des modèles auxquels nous pensons qu'ils se rapprochent le plus. Parfois, ces connaissances peuvent être rattachées à plusieurs modèles. Les résultats complets de cette enquête sont disponibles sur Internet¹³.

3.1 Connaissances pour le modèle de la tâche

Le modèle de la tâche doit permettre au système d'intégrer des notions et des concepts relatifs au thème adressé par l'application. Il s'agit de délimiter l'univers de référence commun aux interlocuteurs.

La Figure 8 ci-après présente la perception qu'ont les utilisateurs du catalogue électronique de la médiathèque disponible sur Internet.

Ceux qui « ne se prononcent pas » n'ont en fait, jamais utilisé ce catalogue. Associés à ceux qui le trouvent insuffisant, cela nous amène à 56% de personnes qui n'utilisent pas cet outil, qui se veut, pourtant, le cœur de la médiathèque.

¹² Remarque d'André Bisseret lors de la présentation des résultats.

¹³ <http://hermes.imag.fr/~rouillar/mediatheque/enquete.html>

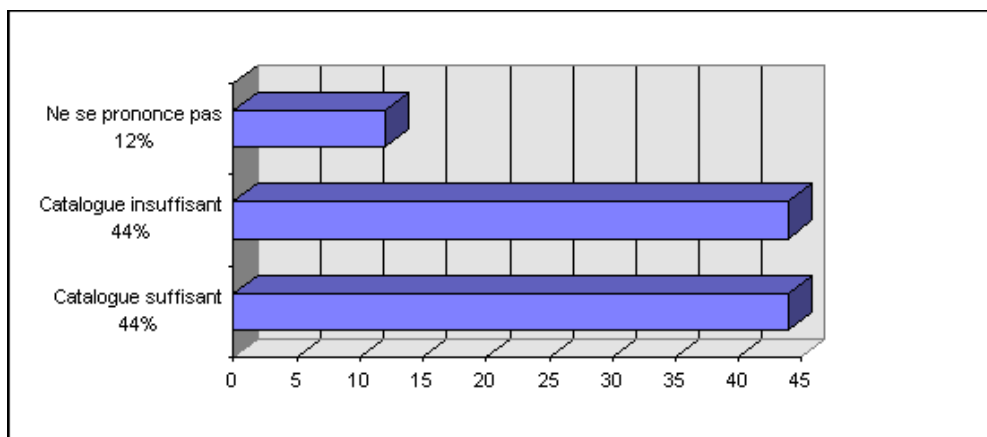


Figure 8 : Avis des usagers sur le catalogue de recherche des documents de la médiathèque

Parmi les réflexions des personnes, voici quelques unes de celles qui reviennent le plus souvent. Elles reprochent au système informatique en place :

- ◆ De ne pas indiquer si le livre est présent (physiquement) à la médiathèque ou s'il a été emprunté.
- ◆ De donner trop peu ou beaucoup trop de détails, sans jamais donner ce qu'ils attendent.
- ◆ De donner des réponses plus longues que pertinentes.
- ◆ De ne posséder qu'un seul champ de sélection, et de ne pas pouvoir ainsi distinguer une recherche par nom d'auteur d'une autre par mot contenu dans le titre (par exemple).
- ◆ De ne pas pouvoir clairement distinguer les ouvrages présents à la médiathèque de ceux d'autres bibliothèques.
- ◆ De ne pas donner assez de détails par rapport à un thème, un domaine précis...
- ◆ Le caractère trop "simpliste" du formulaire.

3.2 Connaissances pour le modèle de l'utilisateur

Le modèle doit permettre au système de s'adapter à son interlocuteur. Dans notre cas, l'usager peut difficilement être remplacé par une autre personne lors de ses recherches. On parle d'utilisateur non générique. Il convient alors de déterminer les principales stratégies qu'utilisent les usagers selon leur catégorie (débutants, confirmés, experts).

Nous avons tout d'abord voulu savoir si les utilisateurs effectuaient des recherches selon un thème ou un domaine précis.

Il semble, selon les résultats obtenus (Cf. Figure 9), que le plus souvent, l'utilisateur sait ce qu'il cherche, mais de manière globale. Il étudie un thème spécifique mais il lui arrive cependant (souvent ou de temps en temps = 68%) d'effectuer des recherches sans vraiment pouvoir définir l'objet précis et détaillé de ses recherches.

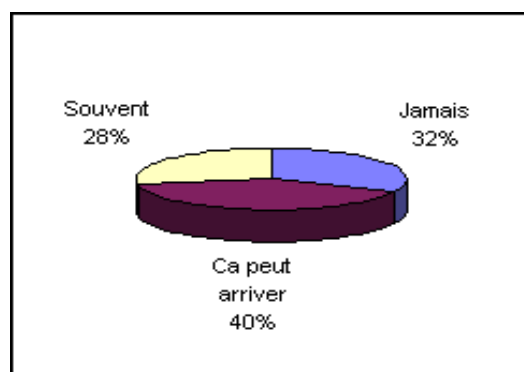


Figure 9 : Recherches selon un thème ou un domaine

Il procède alors par analogie, par glissement vers des thèmes proches ou des domaines frontaliers, pour arriver à ses fins. Ce qui est précis, c'est le thème de la recherche qu'il effectue. Le type de document recherché est également assez précis. En revanche, les caractéristiques mêmes des documents recherchés, comme l'auteur, l'année et/ou le lieu de publication, ne sont pas toujours connues.

La Figure 10 ci-après montre que pour les utilisateurs, la recherche documentaire se fait principalement par tâtonnements 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.

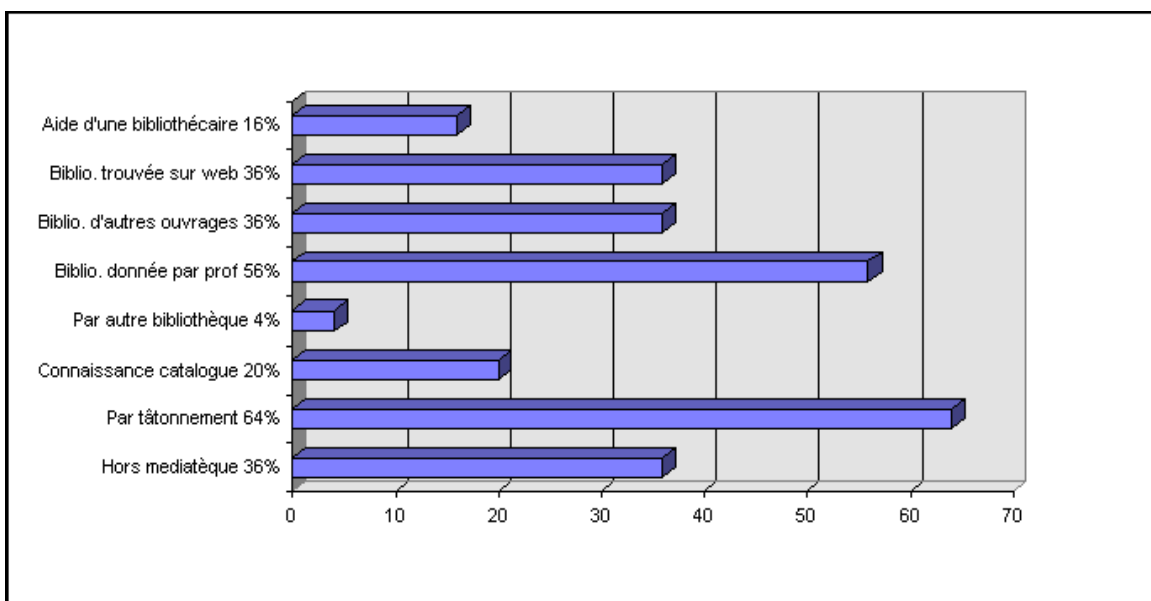


Figure 10 : Modes de recherche de documents des usagers (question à choix multiple)

Cela rejoint les observations de [BEAUFILS 96] qui dit que "*pour atteindre son objectif, le consultant ne sait pas toujours construire un projet de consultation précis et faire évoluer ce projet en fonction des informations qu'il recueille progressivement. Il a souvent du mal à définir le besoin d'informations qui correspond au problème qu'il traite et à identifier dans la base qu'il explore les indices qui peuvent l'amener aux champs d'informations souhaités.*".

Par ailleurs, on observe que 68% des usagers utilisent les terminaux de recherche 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, les 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.

La Figure 11 présente des résultats très intéressants. Ils nous apprennent que les usagers considèrent l'information rapportée par la machine comme étant « plutôt riche », mais ceux-ci s'empressent d'ajouter : « mais pas toujours pertinente ».



Figure 11 : Pertinence des informations données par la machine, selon les utilisateurs

3.3 Connaissances pour le modèle de dialogue

Le modèle de dialogue constitue le cœur d'un système de dialogue. Il doit permettre de gérer les interventions entre les interlocuteurs, et d'identifier leurs types. Il doit pouvoir détecter si le dialogue progresse normalement, si des incidences sont faites, et si des erreurs ou des ambiguïtés interviennent au cours des échanges.

Nous n'avons pas recueilli beaucoup d'informations à propos du modèle de dialogue dans cette étude, mais il est intéressant de noter, sur la Figure 12, la persévérance des usagers lorsque les résultats sont infructueux.

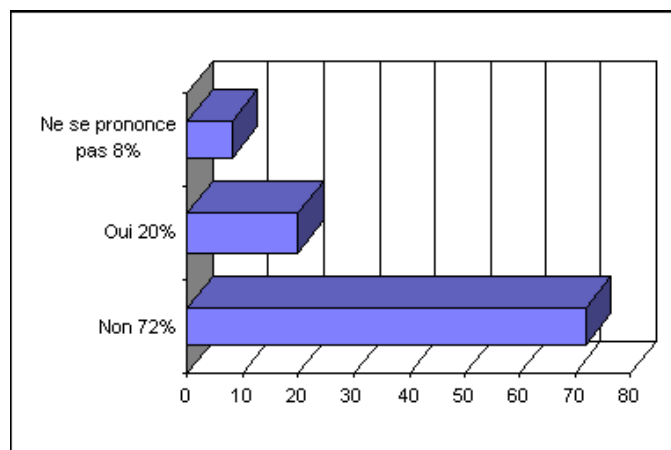


Figure 12 : Interruption volontaire de recherche par les utilisateurs en cas d'incompréhension de la machine

Une très grande majorité des usagers ne renoncent pas lorsqu'ils ne trouvent pas ce qu'ils recherchent, et tentent d'obtenir des informations plus pertinentes en formulant différemment leurs demandes. L'inconvénient majeur constaté est le fait qu'actuellement, l'échange n'est pas interactif et par conséquent qu'aucun dialogue ne peut s'engager entre l'homme et la machine.

En ce qui concerne le DHM plus particulièrement, il semble que, comme le montre la Figure 13, les utilisateurs soient 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.

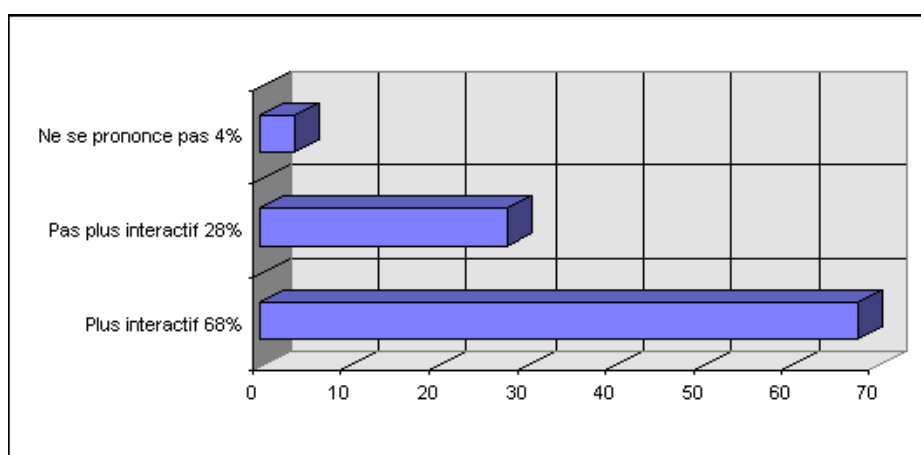


Figure 13 : Souhait d'un dialogue plus interactif avec la machine

Avec la Figure 14, on découvre qu'une part importante des usagers se dit prête à tester une interface vocale pour une recherche documentaire (28 % de *volontiers* + 44 % de *peut-être* =72%), ce qui laisse supposer qu'ils ont besoin d'un meilleur dispositif, permettant une communication Homme-Machine performante et pertinente.

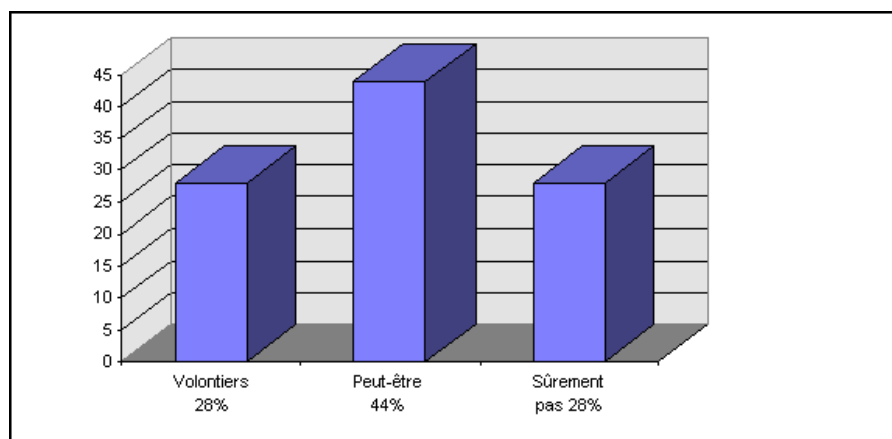


Figure 14 : Hypothèse d'utilisation d'un ordinateur avec reconnaissance de la parole pour la recherche d'ouvrages à la médiathèque ?

Dans l'ensemble, cette enquête révèle que 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.

En résumé :

Cette enquête d'usage révèle que les utilisateurs ont, semble-t-il, des difficultés à interagir avec la machine pour effectuer des recherches d'informations pertinentes, lorsqu'ils utilisent des outils informatiques classiques. L'hypothèse d'un DHM en langue naturelle, voire de l'utilisation de la parole, pour pallier ce problème, est globalement bien acceptée par les personnes interrogées.

Compte tenu des éléments obtenus grâce à cette étude, nous envisageons de concevoir un système de DHM en langue naturelle, ayant les caractéristiques suivantes :

- Collaboratif (mais pas intrusif) pour l'aide à la recherche d'information documentaire,
- Autorisant la consultation d'une base de données structurée,
- Interrogeable à distance via le Web,
- Adaptatif à l'utilisateur (novice ou expert),
- Capable d'initiatives pour orienter le dialogue (affiner, élargir la recherche) en fonction de l'avancement de la tâche, du but et du profil de l'utilisateur,
- Permettant une recherche vague, floue, approximative, lorsque l'utilisateur ne sait pas avec précision ce qu'il recherche.

Ces travaux préliminaires nous donnent des éléments de base et des connaissances initiales pour l'élaboration des modèles de la tâche, de l'utilisateur et du dialogue. Cependant, nous ne disposons pas d'assez d'informations sur le modèle de la langue, pour mettre en œuvre une application qui soit immédiatement utilisable, dans le cadre d'une aide à la recherche d'information documentaire.

Nous allons donc, dans un premier temps, mettre à la disposition des usagers un système informatique, capable de dialoguer à propos d'ouvrages issus d'une base de données accessible sur Internet.

Cela nous permettra de recueillir des informations linguistiques relatives au domaine traité. Nous espérons ainsi obtenir des éléments pertinents pour la définition du modèle de langage, et mettre en œuvre, dans un deuxième temps, un système de DHM plus robuste.

Cette démarche s'inscrit dans notre processus incrémental précité (cf. spirale) comme une étape intermédiaire, mais nécessaire, pour mener à bien notre projet.

« L'ordinateur travaillera main dans la main avec l'homme et exécutera toutes les difficiles besognes qui exigent temps, minutie et précision. La symbiose de l'homme et de l'ordinateur sera, j'en suis certain, un événement fantastique. »

Claude Elwood Shannon, 1938.

4 Mise en œuvre d'un système de DHM

Nous assistons actuellement à l'essor d'applications informatiques interactives touchant à des tâches cognitives complexes, comme la recherche d'information. Dans de telles situations, nous avons vu que l'utilisation de la langue naturelle est potentiellement très intéressante puisqu'elle diminue l'effort cognitif que doit fournir l'utilisateur pour énoncer ses intentions et formuler ses requêtes. Nous pensons donc, comme d'autres chercheurs du domaine, qu'il n'est plus nécessaire, de nos jours, de souligner l'importance d'une interface en langage naturel pour dialoguer avec les bases de données, par exemple ; en particulier dans le cas d'utilisateurs non spécialisés [NOUALI 97].

Quelques systèmes fonctionnent déjà pour des domaines d'application restreints, mais même lorsque les tâches à réaliser par les usagers sont simples, un dialogue langagier convivial et efficace soulève de nombreuses difficultés en raison de l'insuffisance de robustesse des systèmes.

De nombreux travaux relatifs au traitement des bases de données visent à permettre une meilleure communication homme-machine fondée sur un dialogue en LN. Trigano explique que *« les interfaces en langues naturelles améliorent considérablement la communication homme/machine des systèmes à bases de connaissances ainsi que celle des systèmes de gestion de bases de données. Dans le cas de systèmes documentaires ou de systèmes textuels, le langage naturel représente, à notre avis, le type d'interface le mieux adapté. »* [TRIGANO 94].

Cependant, comme le fait remarquer Sabah, à propos du traitement automatique des langues, le problème réside essentiellement aux niveaux sémantique et cognitif : *« actuellement, la sémantique est le goulot d'étranglement pour les mises en œuvre en vraie grandeur, et nous développons l'idée que la prise en considération des aspects perceptuels, en liaison avec les mécanismes d'apprentissage, permettraient une meilleure base sémantique. »* [SABAH 97].

4.1 Problème de circularité pour la mise en œuvre d'un DHM

Pour modéliser les interactions entre l'homme et la machine, les chercheurs s'inspirent principalement d'interactions entre humains. En effet, *« travailler sur un corpus de DHM réel n'est pas envisageable car le système capable de répondre n'existe pas encore. Les limites des systèmes actuels posent des problèmes pour la collecte du corpus. Nous sommes obligée de nous appuyer sur des descriptions du dialogue humain pour modéliser le DHM car c'est tout ce dont nous disposons à l'heure actuelle. »* [TALEB 96].

Pour réaliser des systèmes interactifs, les scientifiques du domaine manquent de corpus de DHM en langue naturelle. 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 [SIROUX et al. 89] : celui de disposer de données pour réaliser un système qui n'a pas encore produit ces données. [POLITY et al. 90] expliquent *« Il y a donc une contradiction : la conception d'un système exige que soit identifié le langage de ses utilisateurs, et cette identification exige l'existence du système. Pour sortir de cette contradiction, la solution fréquemment adoptée est celle d'expérimentation "Magicien d'Oz". »*.

L'insuffisance des connaissances et des modèles actuels constitue l'un des obstacles majeurs des recherches visant à la conception et à la mise en œuvre d'interfaces de dialogue en langue naturelle. Des méthodes ont été mises au point pour atténuer cette carence : les expériences de type Magicien d'Oz et celles de recueil de dialogues par messageries électroniques (type Minitel ou autres). Ces méthodes sont intéressantes. Toutefois, elles peuvent être critiquées et améliorées, comme nous allons le voir ci-après.

4.2 Critique de la méthode Magicien d'Oz

Dans une expérience de type Magicien d'Oz, un compère humain simule les comportements de la machine à l'insu de l'utilisateur. Ce dernier pense donc dialoguer avec une machine, alors qu'en fait, il n'en est rien. Cette technique entraîne quelques contraintes : effort d'organisation, de préparation du matériel, choix de sujets motivés pour se soumettre à l'expérience, prise de rendez-vous, déplacement sur le lieu de l'enregistrement, consignes plus ou moins précises à respecter pour le compère, etc. Ces simulations sont donc délicates et coûteuses à élaborer.

Les corpus de données ainsi recueillis sont souvent peu volumineux, et l'on n'est pas sûr, de surcroît, que les dialogues obtenus dans ces conditions ne soient pas quelque peu biaisés. Il existe, en effet, des comportements langagiers induits par la machine : similitudes, enrichissements, appauvrissements, etc. [LUZZATI 95]. Or, en protocole du Magicien d'Oz, ces phénomènes sont parfois engendrés par les propres comportements du compère qui joue le rôle de la machine.

4.3 Critique de la méthode messagerie électronique

Une autre méthode de recueil de corpus de dialogues en LN consiste à utiliser des messageries électroniques sur lesquelles se connectent des usagers qui posent des questions précises en vue d'obtenir des renseignements. Chanet donne un exemple de

corpus de ce type [CHANET 96] : des étudiants se connectent à une messagerie Minitel et posent des questions diverses sur les cursus, les modalités de recrutement, etc. Mais, de cette manière, on n'obtient que des messages Homme-Homme (novice/expert), et l'on ne peut pas parler de réel dialogue puisque les questions et les réponses sont trop différées dans le temps (environ un jour). De plus, on ne sait pas vraiment qui se connecte, ce qui empêche tout traitement des données selon des catégories distinctes d'utilisateurs.

* * *

Pour progresser vers notre objectif, il est nécessaire de :

- (a) résoudre le problème de la circularité, en développant un système initial puis en l'améliorant progressivement au vu des résultats que l'on obtient en utilisation réelle. Chaque nouvelle version doit permettre d'identifier plus de concepts, et de les intégrer au module de traitement automatique des résultats ;
- (b) de recueillir des corpus via le World Wide Web où le DHM est réel (et crédible), avec un temps de réponse instantané (sauf lenteur du réseau Internet).

Pour ce faire, nous introduisons ci-dessous la notion d'hyperdialogue.

L'être humain appréhende son environnement grâce à ses capacités cognitives et sensori-motrices : vision, ouïe, toucher, odorat, goût, geste, parole. La recherche en communication Homme-Machine s'oriente actuellement vers la conception et le développement de systèmes informatiques coopératifs adaptés aux capacités cognitives des utilisateurs.

Dans cette perspective, nous définissons l'hyperdialogue comme étant un dialogue Homme-Machine, coopératif et finalisé, dans un environnement hypertextuel. L'écrit, l'oral, voire le geste, peuvent être utilisés dans l'hyperdialogue. L'usage des dispositifs est séquentiel ou parallèle, et la circulation de l'information se fait de manière combinée ou indépendante (Cf. [NIGAY et al. 92] et [BOURGUET 92]).

Le cœur du système générique d'hyperdialogue que nous allons présenter réside dans l'interaction coopérative multimodale et finalisée que partagent la machine et l'utilisateur.

Pour nous, le dialogue est un processus tendant à maximiser l'accord entre les interlocuteurs à propos d'une tâche qu'ils ont à faire en commun. Différentes stratégies de dialogue doivent être envisagées, selon les types d'utilisateurs 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**, les interlocuteurs coopèrent grâce au dialogue, afin que l'objet de la requête soit clair et précis. En stratégie **réactive**, la réponse à une question est apportée dès que possible, quitte à faire des hypothèses s'il existe des ambiguïtés dans l'énoncé. 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 agents s'éloignent trop l'un de l'autre. Les interlocuteurs vont devoir négocier, chacun d'eux essayant d'obtenir une solution optimale.

4.4 Système générique d'hyperdialogue

Dans la Figure 15, ci-après, nous présentons un système générique d'hyperdialogue dédié au WWW :

- Les éléments entrants peuvent être de plusieurs types : traditionnel (clavier + souris), verbal ou gestuel.
- Le module de reconnaissance de parole fournit en sortie une chaîne de caractères pour le module d'interprétation linguistique.
- Le processus d'interprétation gestuelle peut fournir des coordonnées et des mesures de déplacement.
- Les événements liés à la souris sont recueillis de manière traditionnelle (événement système 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.

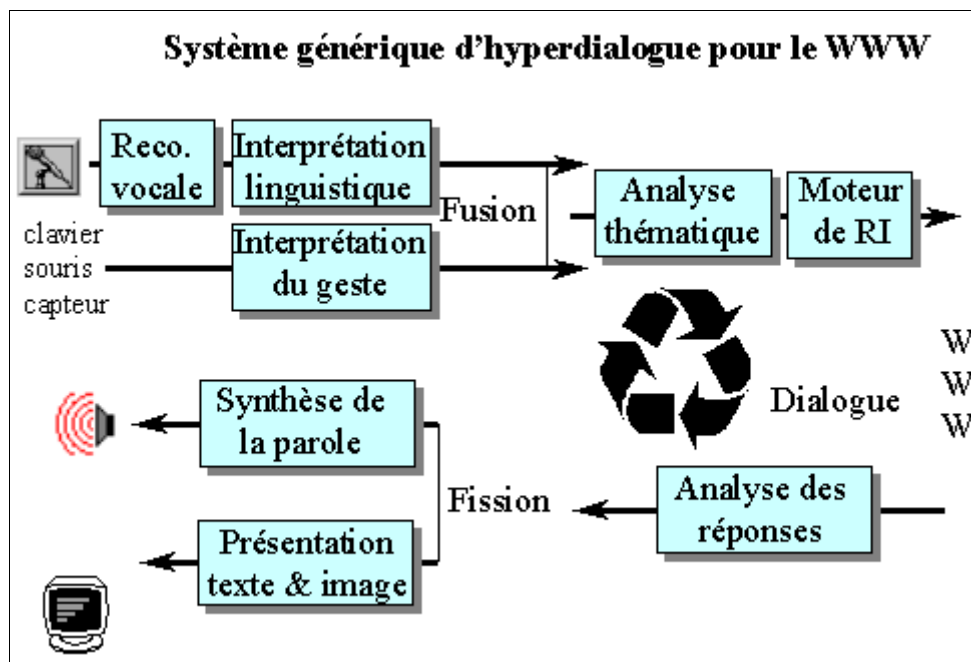


Figure 15 : Un système générique d'hyperdialogue pour le WWW

4.5 Apports de l'hyperdialogue

Nous considérons que l'apport de l'hyperdialogue est principalement de deux ordres. D'une part il peut contribuer au développement des systèmes de DHM en soi, en utilisant un support de communication encore peu exploité jusqu'à présent, et d'autre part, il donne des éléments intéressants pour l'étude scientifique du DHM.

4.5.1 Apports aux systèmes de DHM

Un système d'hyperdialogue favorise la convergence entre un DHM dit classique, et une navigation dans un environnement hypertextuel, comme sur le Web. Le fait de coupler navigation et DHM en LN permet de donner de l'information relativement au contexte. On imagine, par exemple, un hyperdialogue sur le site d'un hôtel présentant des images des chambres et services proposés. À la question « combien mesure-t-elle ? », le système saurait donner la superficie de la chambre en question. Il serait également capable de répondre à la même question, mais en présence de la photo de la piscine de l'établissement, cette fois-ci.

En navigation hypertextuelle au sein de résultats, il sera utile de cliquer sur un hyperlien (du texte, mais aussi des photos ou des vidéos) pour obtenir des informations intéressantes, en navigant de liens en liens. Le dialogue sera à son tour efficace pour obtenir des informations complémentaires qu'il n'est pas possible d'obtenir en manipulation directe au clavier/souris, et pour alléger la charge cognitive de l'utilisateur, qui pourra exprimer ses besoins en langue naturelle, écrite ou orale.

4.5.2 Apports à l'étude du DHM

Pour l'étude du DHM, les avantages d'une telle approche sont les suivants. Le système est disponible 24h/24, accessible de n'importe où dans le monde, utilisable par tous. Il est possible d'utiliser (en ligne) des ressources externes que le système n'a pas à supporter (un dictionnaire, une banque d'images, etc.).

Scientifiquement, il est intéressant de noter que le système sort du cadre du laboratoire, pour s'ouvrir sur le monde réel. Il est mis à la disposition d'un large public, hétérogène, et peut être critiqué, en temps réel, au fur et à mesure des améliorations proposées. Les corpus recueillis directement au format électronique (pas de retranscription) sont volumineux, et le dépouillement peut aisément faire l'objet d'un traitement automatisé.

En résumé :

Nous avons discuté le problème de circularité qui empêche, pour un système que l'on veut mettre en œuvre, de s'appuyer sur les connaissances que pourraient apporter les corpus recueillis par ce propre système, puisque ce dernier n'existe pas encore.

La modélisation d'un DHM coopératif et finalisé, ainsi que la réalisation d'un système complet, peuvent, à notre sens, s'inscrire dans un cycle de développement incrémental, sans utiliser de techniques du type Magicien d'Oz ou messagerie électronique, que nous avons critiquées.

La notion d'hyperdialogue que nous venons de présenter nous semble intéressante pour coupler un DHM avec une navigation hypertextuelle. Une telle approche peut faire évoluer la discipline, notamment grâce au recueil de corpus de véritables dialogues entre l'homme et la machine, non biaisés par les comportements d'un expérimentateur (qui joue le rôle de la machine).

Pour exploiter pleinement les possibilités offertes par le support de communication choisi (le Web), nous devons ajouter aux caractéristiques du système que nous voulons mettre en œuvre, la gestion d'une multimodalité exclusive (interaction au clavier ou bien à la voix, par exemple), et offrir à l'utilisateur la possibilité d'agir sur l'interface du système grâce au dialogue en LN.

Le système générique d'hyperdialogue présenté, doit servir de référence pour la mise en place de notre système de DHM multimodal sur le Web. Nous espérons ainsi recueillir un corpus de dialogues original et exploitable, pour en extraire, entre autres, les données nécessaires à l'élaboration de notre modèle de la langue.

5 Le système Halpin-Recueil

Que faut-il entendre par DHM en langue naturelle ? 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. Le système HALPIN fonctionne par reconnaissance de concepts dans un dialogue coopératif et finalisé entre un humain et une machine [ROUILLARD 98].

Pour la modélisation du DHM nous sommes face à deux problèmes principaux :

- 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, nous avons procédé à une simulation de la tâche de recherche d'information documentaire assistée par un vrai système de dialogue ayant une stratégie coopérative (nous décrivons cette expérience ci-après). Cela permet de capturer les représentations conceptuelles du sujet et d'y associer les énoncés utilisés. Puis nous procédons à 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 réduisons l'analyse des énoncés à une analyse conceptuelle afin de modéliser le niveau de compréhension du système. Il est possible alors d'incrémenter 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.

Ce premier système a pour vocation de recueillir des informations permettant de mieux modéliser le DHM. Notre objectif n'est pas de comparer les résultats que l'on va obtenir avec des dialogues Homme-Homme, pour voir si la machine est plus ou moins efficace, si le but est effectivement atteint ou pas, etc. Nous espérons simplement recueillir un corpus de dialogues réels entre un homme et une machine, afin de disposer de données propices à une réflexion autour d'une modélisation du DHM via le Web.

5.1 Principe de l'expérience de recueil de dialogues

Durant cette expérience sur le WWW, l'utilisateur participe à une sorte de jeu interactif avec la machine. Il se connecte à une page Web de notre serveur, qui lui présente la tâche à accomplir. L'ordinateur choisit un livre au hasard, parmi 39931 documents disponibles, dans une bibliothèque électronique universitaire grenobloise¹⁴ (l'université Pierre Mendès France, spécialisée en sciences humaines et sociales), et en mémorise les principales caractéristiques (auteur, éditeur, mots-clés, etc.). Le but fixé à l'utilisateur est de retrouver l'ouvrage choisi au hasard, en posant à l'ordinateur des questions en LN. La machine tente de comprendre les énoncés et d'y répondre correctement. L'utilisateur a la possibilité d'interagir avec le système en manipulant les résultats obtenus lors de requêtes. Il peut décider de lancer une requête ou bien de continuer à dialoguer pour affiner ses connaissances. L'objectif est atteint lorsque l'utilisateur donne à la machine le numéro ISBN (International Standard Book Number) correspondant au livre en question.

5.1.1 Intérêt de cette méthode

Il s'agit d'une recherche documentaire particulière, en ce sens que l'on demande à l'utilisateur de retrouver un livre sur lequel il ne détient au départ aucune information. A l'inverse de ce qui existe traditionnellement, c'est l'humain qui doit découvrir une

¹⁴ <http://dodge.upmf-grenoble.fr:8001/fra/catalogues.html>

information que possède la machine. Cette situation est bien peu courante en réalité, mais ici, le prétexte du jeu va permettre :

- d’apprécier la validité de notre méthode de recueil de corpus via le Web,
- d’analyser les représentations mentales que se font les personnes à la recherche d’un ouvrage (ses caractéristiques),
- d’étudier les stratégies de dialogue employées, la logique d’enchaînement des tours de parole et la composition linguistique des énoncés (le type de langage employé lors d’un dialogue avec une machine),
- de déterminer de quelle manière les usagers conçoivent un dialogue en langue naturelle (que s’autorisent à dire les usagers en LN face à une machine ?).

5.1.2 Modalités techniques

Techniquement, chaque énoncé (de l’homme et de la machine) est sauvegardé et daté (heure, minute, seconde) dans un fichier de dialogue. La phrase de l’usager est envoyée via le Web à un outil de lemmatisation développé par Xérox [GAUSSIER et al. 97]. Cela permet de travailler sur la forme canonique de la phrase. On ne traite alors qu’une forme simplifiée de l’énoncé (sans conjugaison, sans formes fléchies, sans pluriel, etc.). Les éléments qui composent cette phrase canonique sont comparés à une liste de concepts et en cas de correspondance, des informations sont données par le système.

La démarche utilisée est incrémentale. Nous sommes ici dans l’une des parties de la spirale de développement. Un vocabulaire de base a été préparé, au regard d’autres travaux en DHM [MOREL 88], [BILANGE 92]. Puis la première version a été annoncée sur les listes de diffusion électroniques et des utilisateurs volontaires de tous types ont pu interagir avec le système, tester ses limites, et le critiquer pour soulever des points à améliorer. À chaque nouvelle analyse des résultats obtenus, nous affinons la pertinence du système de dialogue, en lui ajoutant des mots ou des concepts qu’il ne savait pas traiter jusque là. La base de concepts est indépendante de tout code logiciel et se compose uniquement de données au format texte. Ainsi, une simple mise à jour dans le fichier de concepts permet d’augmenter les « connaissances » du système, sans avoir à recompiler quoi que ce soit.

Avec ce système évolutif et notre approche de dialogue par reconnaissance de concepts, nous rejoignons des travaux comme ceux de [PONTON 96] ou de [LEHUEN 97], qui tendent à définir un système noyau, non dédié, autour duquel on apporte des éléments de vocabulaire spécifiques à une tâche précise.

L'effort a porté, d'une part, sur la conception d'un modèle permettant l'identification et la reconnaissance de concepts dans le dialogue, et d'autre part, sur la réalisation d'un système fonctionnant sur le réseau Internet et pouvant être utilisé par un grand nombre d'utilisateurs.

Le système ainsi réalisé se base sur la notion d'hyperdialogue et vise à récolter des dialogues réels entre un humain et une machine via le WWW. L'utilisateur demeure anonyme et reste libre de sa stratégie de recherche. Contrairement aux expériences en protocole du Magicien d'Oz, le sujet, sachant qu'il n'est pas directement observé, peut sans doute se permettre d'agir plus naturellement qu'il ne l'aurait fait en laboratoire, avec des caméras et des personnes qui notent le moindre de ses faits et gestes. On observe par exemple dans notre corpus des plaisanteries ou des insultes que l'on observe peu dans des corpus traditionnels.

L'architecture que nous présentons ci-dessous permet un emploi du système par plusieurs utilisateurs simultanément, vingt-quatre heures sur vingt-quatre. Nous avons ainsi pu obtenir plus d'un millier de dialogues réels grâce à cette méthode. Seuls 980 fichiers 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.

5.2 Architecture du système Halpin-Recueil

Pour cette méthode de recueil de corpus, nous utilisons le World Wide Web comme support matériel. Le système Halpin a été développé en langage Perl et Html. La Figure 16, ci-après, présente l'architecture du système. L'utilisateur se connecte à la page Web Halpin, du serveur du laboratoire CLIPS à Grenoble, qui lui présente la tâche à accomplir, de manière textuelle et vocale (un fichier sonore est encapsulé dans la page

d'accueil). Un ouvrage est choisi de manière aléatoire sur le serveur de la bibliothèque, et le dialogue entre l'homme et la machine débute, à l'initiative de l'utilisateur. Pour comprendre ce qui a été dit, on utilise un outil morphologique d'analyse de phrase, disponible sur le serveur Web de Xérox.

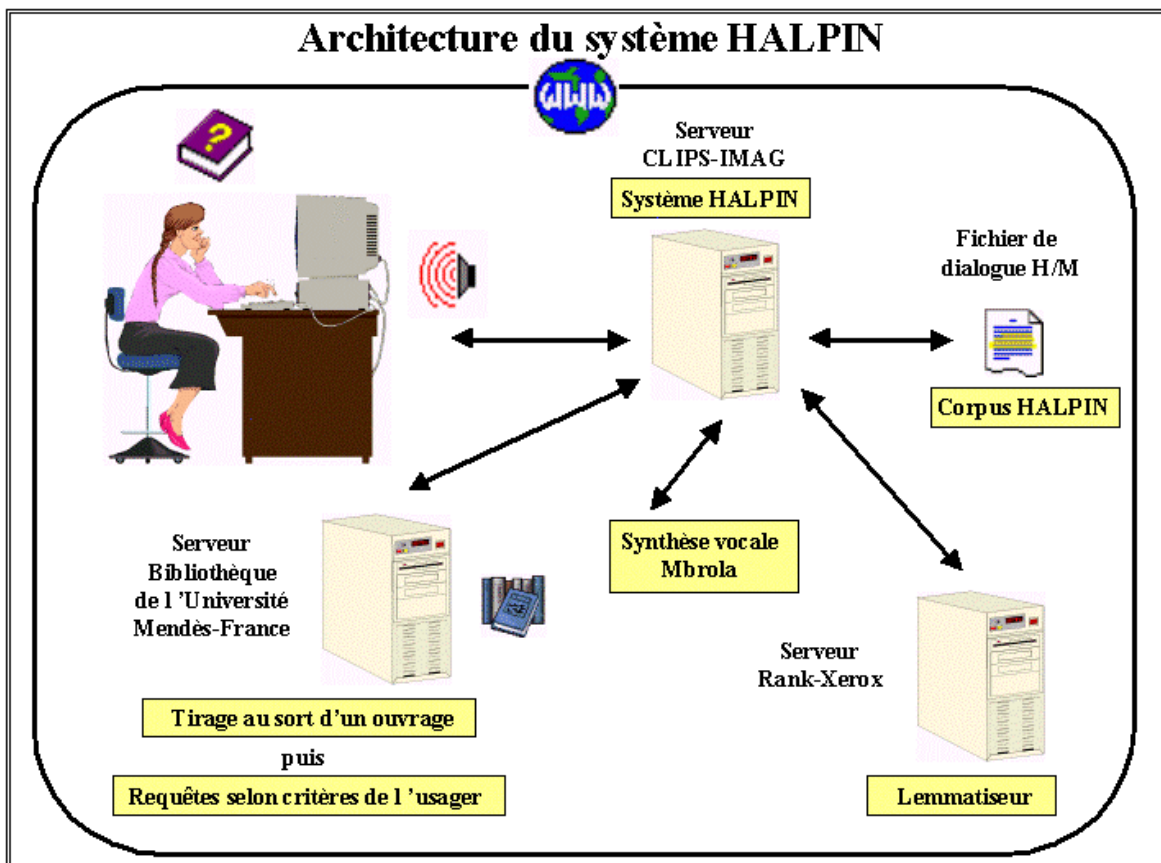


Figure 16 : Architecture du système Halpin-Recueil

Après la phase d'envoi de la requête au lemmatiseur de Xérox, via Internet, le résultat obtenu est sauvegardé dans un tableau à trois colonnes : mot, lemme, catégorie. Le Tableau 1 ci-dessous montre le résultat d'une lemmatisation pour la phrase : " de quoi parle ce livre ? ". On obtient ainsi la forme canonique de l'énoncé de l'utilisateur.

MOT	LEMME	CATÉGORIE
de	de	+PREP_DE
quoi	quoi	+RELPRO
parle	parler	+VERB_P3SG
le	le	+DET_SG
livre	livre	+NOUN_SG
?	?	+SENT

Tableau 1 : Exemple de lemmatisation de « de quoi parle le livre ? »

Le Tableau 2 ci-dessous présente le résultat de la lemmatisation de la phrase " quelle est l'annee, alors ? ". On observe que l'outil de lemmatisation utilisé rétablit les accentuations omises. Ainsi, pour le mot " annee " (sans accent) tapé par l'utilisateur, le lemme " année " sera retourné.

MOT	LEMME	CATÉGORIE
quelle	quel	+RELPRO
est	être	+VAUX_P3SG
l'	le	+DET_SG
annee	année	+NOUN_SG
,	,	+CM
alors	alors	+ADV
?	?	+SENT

Tableau 2 : Exemple de lemmatisation de « quelle est l'annee, alors ? »

Ensuite, une requête est créée par la machine selon les critères identifiés par l'utilisateur, et la base de données de la bibliothèque est interrogée. Les résultats sont affichés à l'écran sous forme d'hyperliens consultables par simple clic souris. Une fiche signalétique (une notice) donne des informations pour chacun des ouvrages de la liste. Le dialogue peut alors continuer. L'utilisateur affine ses requêtes jusqu'à trouver le livre en question et propose à la machine l'ISBN qu'il croit être celui qu'il fallait retrouver, ou bien abandonne s'il ne trouve pas l'expérience intéressante.

A chaque nouvelle connexion d'un utilisateur, le système détermine un numéro unique (processus Unix) pour pouvoir sauvegarder le fichier de dialogue. Puisque l'on bénéficie de l'architecture clients-serveur du réseau Internet, il est pertinent de remarquer que de multiples connexions peuvent se faire simultanément, à n'importe quel moment du jour ou de la nuit.

5.3 Fonctionnement du système d'interaction

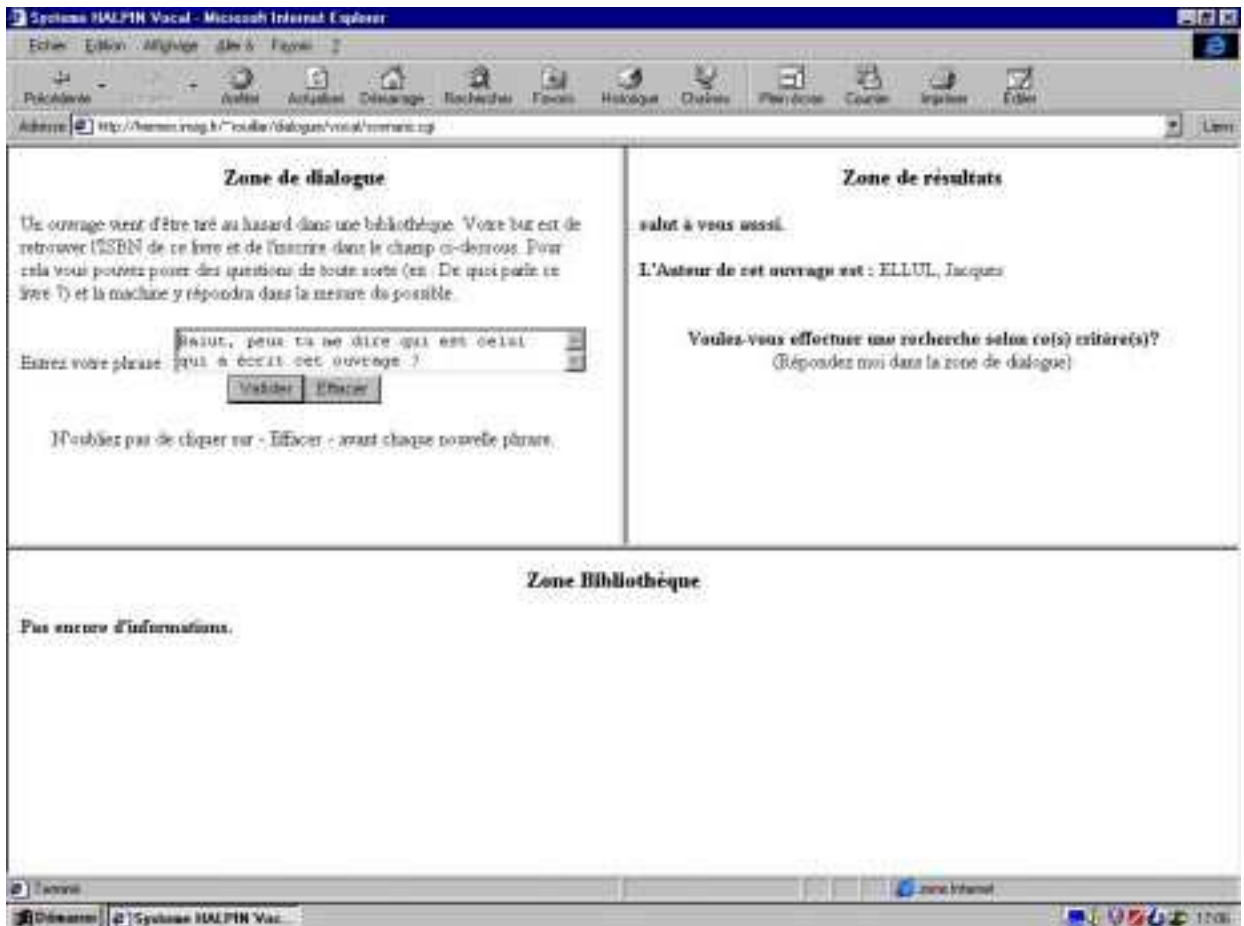
Voici à présent un aperçu du mode de fonctionnement du système Halpin-Recueil. Le système analyse l'énoncé de l'utilisateur et lui offre son aide en proposant de créer une requête afin de rechercher des informations selon le ou les critères qu'a mentionné la personne.

Comme le montre la Capture 1, des cadres (*frames*) sont définis pour segmenter l'écran en trois zones distinctes (zone de dialogue, zone de résultats et zone bibliothèque).

Une fonction de tirage aléatoire est exécutée pour choisir un ouvrage dans une bibliothèque accessible via le Web. Concrètement, cela se fait avec l'appel suivant :

```
$url = new URI::URL("http://dodge.upmf-grenoble.fr:8001/cgi-bin/webbdr/recup?base=soc&ind=$numero");  
my $request = new HTTP::Request('GET', $url);
```

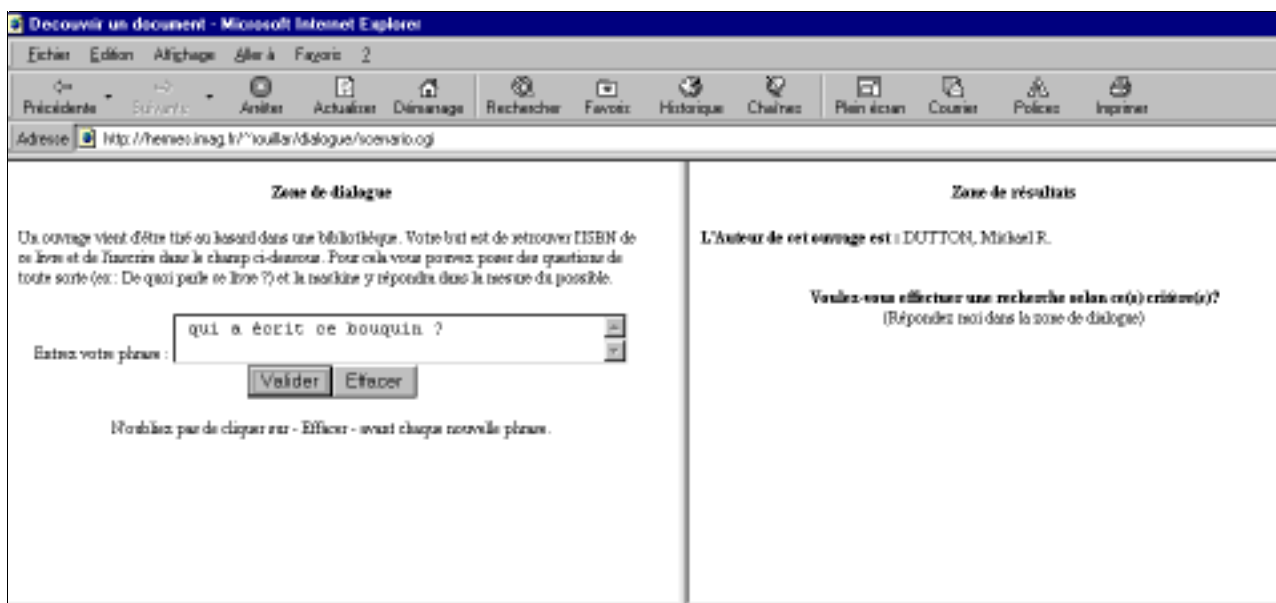
L'argument "base=soc" indique au serveur que la base de données consultée est celle de sociologie, et la variable "\$numero" est remplacée par la valeur tirée au hasard. Les messages sonores d'accueil sur le système Halpin sont statiques. Ils ont été calculés une fois pour toutes, stockés dans des fichiers (dont le poids est fixe), et le navigateur se contente de les appeler comme une ressource traditionnelle (une image par exemple).



Capture 1 : Interface du système Halpin-Recueil

En revanche, dans la suite des interactions Homme-Machine, et puisque l'on ne sait pas à l'avance quel livre sera tiré au sort par la machine, les réponses formulées doivent être calculées " à la volée ".

Étant sur Internet, en architecture client-serveur, il faut effectuer ces opérations pour plusieurs utilisateurs potentiels simultanés : la synthèse vocale calculée en réponse à l'énoncé d'un usager ne doit pas être jouée pour un autre.

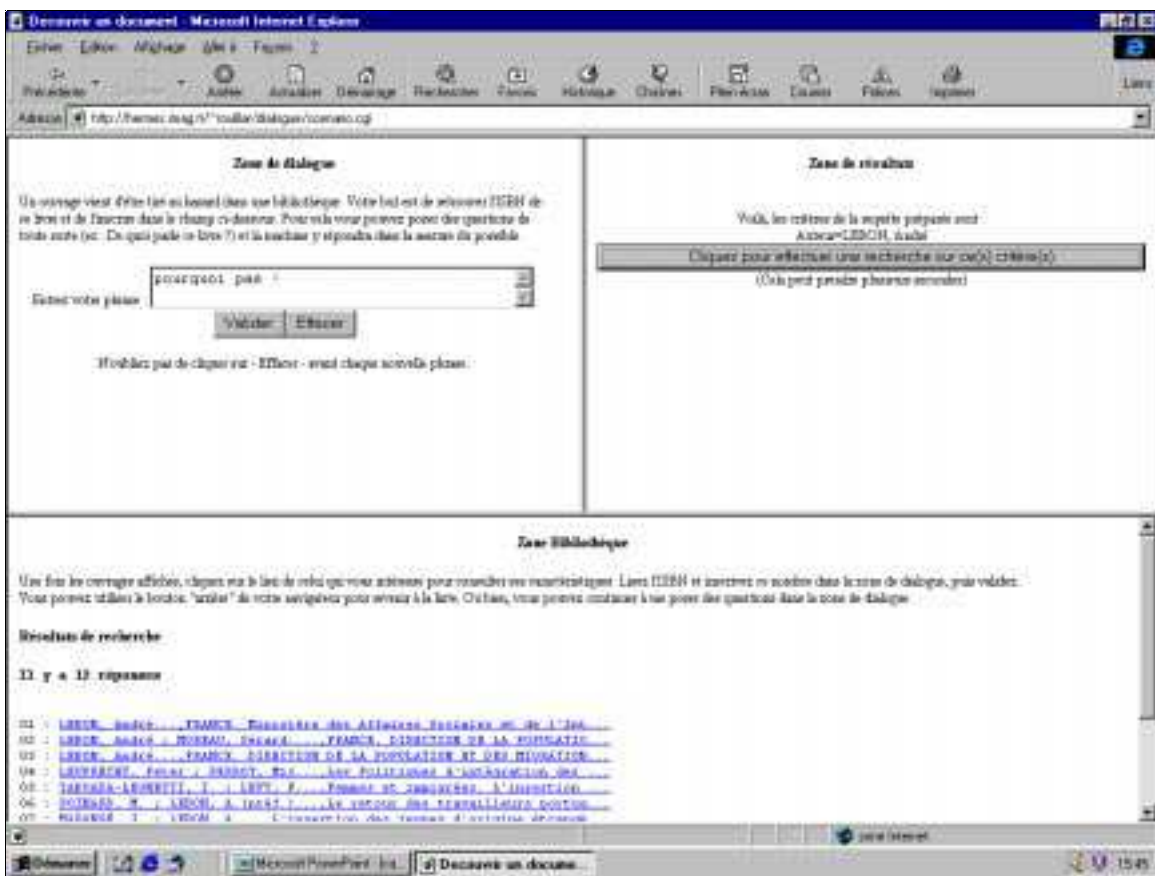


Capture 2 : Interface du système Halpin-Recueil sur le Web : l'utilisateur a posé une question dans la zone de dialogue et la machine a répondu dans la zone de résultats

Dans notre exemple (cf. Capture 2) l'utilisateur demande, dans la boîte de dialogue : « qui a écrit ce bouquin ? ». Le système utilise un algorithme de reconnaissance de concepts (que nous détaillerons plus loin), plus puissant qu'une simple détection de mots-clés. En effet, ici, la machine comprend que l'homme cherche à obtenir des renseignements sur l'auteur, alors même que le mot clé « auteur » n'est pas mentionné. De manière coopérative, après avoir donné l'information désirée dans la zone de résultats de manière textuelle et vocale (« L'auteur de cet ouvrage est : Dutton, Michael R. »), la machine demande à l'utilisateur s'il souhaite effectuer une recherche dans la bibliothèque sur le critère qui vient d'être trouvé (l'auteur). L'homme répond en LN et poursuit ainsi le dialogue avec la machine.

Dans la Capture 3, on observe, avec un autre exemple, le fonctionnement de l'hyperdialogue entre l'homme et la machine : l'utilisateur accepte de faire une recherche en formulant sa réponse en zone de dialogue (il énonce " pourquoi pas ! "), et la machine prépare une requête selon le critère adéquat (dans notre exemple, il s'agit de l'auteur).

En validant cette requête depuis la zone de résultats grâce à un clic souris sur le bouton gris, l'utilisateur accède à la liste des réponses hypertextuelles qu'il pourra consulter en zone bibliothèque. Il sera donc en phase de navigation dans la liste de liens hypertextes (12 réponses avec comme auteur « Lebon André » dans notre exemple). Un clic souris sur une ancre de la liste, dans la zone de bibliothèque, permet d'obtenir des renseignements complémentaires (date de publication, éditeur, ISBN, etc.) sur ce document. En revenant au dialogue, l'utilisateur pourra soit proposer un ISBN comme solution, soit affiner sa requête derechef, grâce au dialogue.



Capture 3 : Hyperdialogue entre l'homme et la machine

5.4 Résultats pour le système Halpin-Recueil

Cette partie présente les principaux résultats obtenus avec le système Halpin-Recueil, en utilisation réelle auprès des usagers du Web. Nous présentons tout d'abord des exemples de dialogues, puis nous montrons en quoi ces résultats sont similaires ou différents des ceux obtenus de manière classique, avec d'autres systèmes de DHM, c'est-à-dire, sans utilisation d'Internet.

5.4.1 Présentation des résultats

Grâce aux informations obtenues lors liaisons entre un client et notre serveur, nous avons pu identifier la provenance territoriale des utilisateurs et dresser une carte des connexions. La Figure 17 montre l'origine géographique des sujets s'étant connectés au système, lors du premier mois d'annonce (grâce à des listes de diffusion, des messages dans les « news » d'Internet, par courrier électronique, etc.).

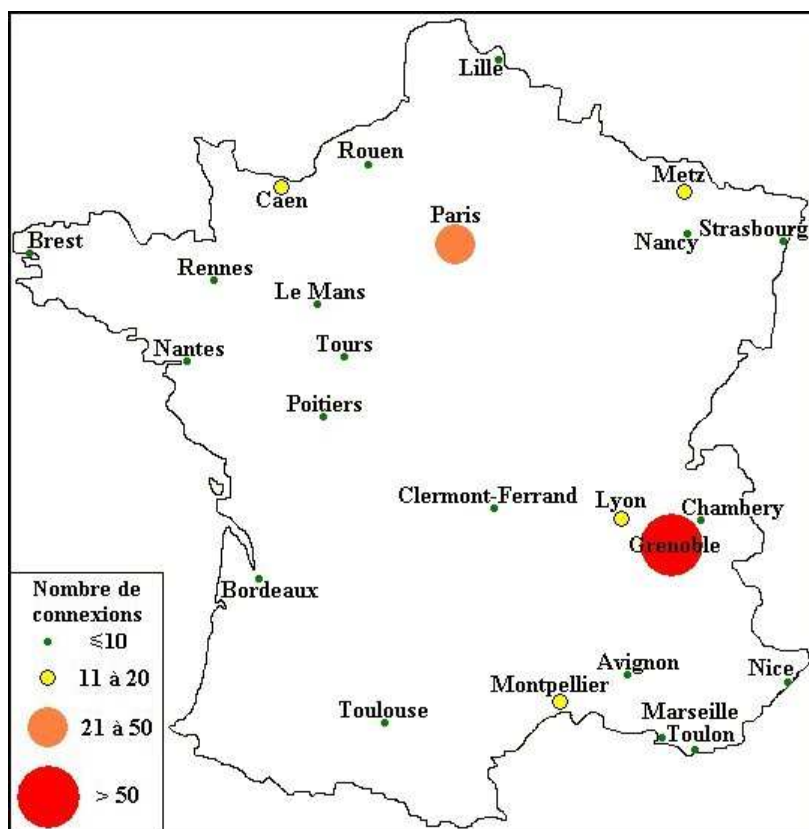


Figure 17 : Carte de France des connexions au système Halpin-Recueil

Au fur et à mesure des connexions, on arrive à obtenir une large couverture géographique de la France métropolitaine, chose qui serait évidemment plus difficile lors d'expériences classiques en

laboratoire, où ce sont principalement des sujets locaux qui participent aux campagnes d'expérimentation.

Ces villes sont toutes universitaires, et il est vrai que ce sont probablement des étudiants ou des chercheurs qui forment la plus grande partie des personnes contactées. Quelques villes relativement proches ont été regroupées sur la représentation graphique. De plus, cette carte ne rend compte que des sites dont le nom est explicite et rapidement identifiable (ex: cormoran.lannion.cnet.fr), et non pas des adresses IP qu'il faudrait tracer pour retrouver la ville d'origine de la connexion (ex : 194.57.187.213).

Il y a eu également bon nombre de connexions depuis l'étranger : Suisse, Québec, Belgique, Australie... Ce sont principalement des pays où l'on parle français. Mais on a également parfois affaire à des utilisateurs français expatriés dans d'autres régions du monde, non francophones.

Le système Halpin a été testé, dans le cadre du projet Orion, sur une liaison entre Grenoble et Lyon. Le délai était d'environ 6 secondes d'attente, en moyenne, entre le moment de la validation d'un énoncé et celui de la prononciation de la synthèse de la réponse. La taille des fichiers sonores joués est variable. En moyenne, elle est de l'ordre de 100 Ko, pour environ 6 secondes d'énoncés (ce qui représente à peu près 2 phrases de 8 mots chacune).

5.4.1.1 Exemples de données recueillies avec le système Halpin-Recueil

Voici plusieurs exemples de fichiers de DHM obtenus grâce au système Halpin-Recueil. Les deux premiers sont des exemples dans lesquels l'utilisateur arrive à accomplir la tâche demandée, et les deux suivants montrent les limites du système lors de cette première version. Une ligne d'en-tête de fichier contient des informations relatives à la machine cliente utilisée par l'utilisateur lors de l'expérience, et indique la date de connexion. Chaque énoncé est daté (heure, minute, seconde). Par convention, « Q » indique une question (de l'utilisateur) et « R » une réponse (de la machine) même s'il est évident que Q et R ne seront pas forcément des questions et des réponses au sens strict (exemple : Q= Bonjour).

niepce.u-grenoble3.fr LE Wed Dec 10 19:02:04 MET 1997

19:02:45:Q= quel est le sujet du livre

19:02:49:R=Les mots-clefs de ce livre sont : Criminologie; Pénologie; Contrôle social; Administration pénitentiaire; Centre spécialisé; Peine; Peine de substitution

19:02:49:R=Voulez-vous effectuer une recherche selon ce(s) critère(s)?

19:03:08:Q=oui

19:04:48:Q=le sujet est-il pénologie ?

19:04:52:R=Les mots-clefs de ce livre sont : Criminologie; Pénologie; Contrôle social; Administration pénitentiaire; Centre spécialisé; Peine; Peine de substitution

19:04:52:R=Voulez-vous effectuer une recherche selon ce(s) critère(s)?

19:05:23:Q=non
 19:05:23:R=D'accord, on ne fait pas de recherche pour l'instant.
 19:05:28:R=Dialoguons ensemble encore un peu...
 19:05:52:Q=quel est l'auteur du livre ?
 19:05:56:R=L'Auteur de cet ouvrage est : FOUCAULT, Michel
 19:05:56:R=Voulez-vous effectuer une recherche selon ce(s) critère(s)?
 19:06:10:Q=oui
 19:08:10:Q=le titre du livre est-il en français ?
 19:08:16:R=Le titre de ce livre est : Alternatives à la prison : Diffusion ou décroissance du contrôle social
 19:08:16:R=Je ne possède pas d'indications à propos de la langue du livre, mais le titre vous donnera cette information là.
 19:08:16:R=Dialoguons ensemble encore un peu...
 19:08:57:Q=2760624420
 19:09:01:R=BRAVO, vous avez retrouvé le bon livre !

Capture 4 : Un exemple de dialogue avec le système Halpin-Recueil

Dans l'exemple de la Capture 4, le dialogue s'articule autour de huit tours de parole. Notons que le dialogue est présenté tel quel, c'est-à-dire sans correction des erreurs de frappe ou d'accentuation de l'utilisateur. La machine donne des informations relativement pertinentes par rapport aux énoncés de l'utilisateur, et cette coopération permet de retrouver le bon numéro ISBN du livre qu'il fallait identifier. Mais on voit bien que les répliques de l'ordinateur sont largement perfectibles. Par exemple, à la question «19:04:48:Q=le sujet est-il pénologie ? », il serait plus judicieux que la machine réponde « oui » plutôt que de répéter la liste des mots-clés. De même lorsqu'on lui demande « 19:08:10:Q=le titre du livre est-il en français ? », il serait plus pertinent de répondre « oui » ou « non » ...

Mais comme l'aura compris le lecteur, notre but n'était pas de rendre ce premier système Halpin optimum, mais bien de collecter des données issues d'un véritable DHM en vue de mettre au point un deuxième système, plus élaboré, et surtout, faisant réellement office d'assistant pour une recherche documentaire.

Dans l'exemple suivant (Capture 5), on observe, comme nous l'avions annoncé plus haut, que les utilisateurs semblent dialoguer sans contraintes avec la machine, et utilisent parfois un langage familier, ou bien tentent de plaisanter pour tester les capacités de la machine.

leucippe LE Fri Feb 20 16:51:36 MET 1998
 16:55:25:Q= qui est l'auteur
 16:55:30:R=L'Auteur de cet ouvrage est : CROFT, John ; SIMPSON, Oliver.(Préf.)
 16:55:30:R=Voulez-vous effectuer une recherche selon ce(s) critère(s)?
 17:07:55:Q=c'est bien ca mon grand

17:08:00:R=Je ne comprends pas le sens de : c'est bien ca mon grand
 18:57:37:Q=t'es nul alors ?
 18:57:37:R=C'est un dialogue en langage naturel, pas un concours d'insultes.
 18:57:50:R=Dialoguons ensemble encore un peu...
 18:59:51:Q=ok, donne moi l'auteur, l'annee, la collection
 18:59:56:R=L'Auteur de cet ouvrage est : CROFT, John ; SIMPSON, Oliver.(Préf.)
 18:59:56:R=La date de publication est : 1980.
 18:59:56:R=La collection est : Home Office Reasearch Study ; vol. 59
 18:59:56:R=Voulez-vous effectuer une recherche selon ce(s) critère(s)?
 19:00:19:Q=oui
 19:02:30:Q=0113406991
 19:02:35:R=BRAVO, vous avez retrouvé le bon livre !

Capture 5 : Exemple de dialogue avec plaisanterie

Au passage, notons l'ambiguïté qui intervient lorsque la machine doit interpréter « 17:07:55:Q=c'est bien ca mon grand ». Ca pourrait être un acquiescement pour répondre à la question posée par la machine (que l'on transformerait en « Oui, c'est bien cela que je veux faire »), ou bien, sur un autre ton, ca pourrait être une sorte de compliment (ironique ?) sur le travail accompli par la machine. Dans l'hypothèse d'une interaction vocale, une étude de la prosodie de l'énoncé donnerait sans doute des informations pour désambigüiser cette phrase [MELONI et LANGLAIS 96]. Ici, d'une part l'énoncé était écrit, et d'autre part, le système ne savait pas que « ce être bien cela » était équivalent à un accord de l'utilisateur. La machine explique donc qu'elle n'a pas compris cette intervention. Le dialogue présenté dans la Capture 6 ci-dessous montre comment un utilisateur se désintéresse d'un dialogue avec la machine lorsque celui-ci glisse vers l'axe incident et que l'ordinateur ne semble pas comprendre ce qu'on lui demande. Le manque de compréhension de la machine va vite entraîner un abandon de l'utilisateur, puisqu'il ne trouve pas un interlocuteur suffisamment performant pour l'aider à accomplir sa tâche. Cependant, même si l'échec de la tâche à accomplir est flagrant, d'un point de vue dialogique, les réponses de la machine restent cohérentes.

gumby.inrs-telecom.quebec.ca LE Wed Dec 10 20:02:09 MET 1997
 20:03:38:Q=qu'y a-t-il marqué sur la page de garde ?
 20:03:42:R=Je n'ai absolument aucune information à propos du nombre de pages de ce livre.
 20:05:29:Q=puis-je voir la page de garde ?
 20:05:33:R=Je n'ai absolument aucune information à propos du nombre de pages de ce livre.
 20:06:20:Q=alors tant pis
 20:06:25:R=Je ne comprends pas le sens de : alors tant pis
 20:07:06:Q=bye
 20:07:10:R=Au revoir, et merci d'avoir participé.

Capture 6 : Un exemple d'échec d'aide à la résolution de la tâche

A l'opposé, dans l'exemple de la Capture 7 ci-après, on voit bien que c'est par manque de robustesse de la part de la machine que le dialogue s'enlise.

cormoran.lannion.cnet.fr LE Thu Jan 15 16:44:21 MET 1998

(...)¹⁵

16:53:57:R=L'Auteur de cet ouvrage est : PRINA, Franco

16:53:58:R=Voulez-vous effectuer une recherche selon ce(s) critère(s)?

16:54:18:Q= je l'ai déjà demandé.

16:54:23:R=Je ne comprends pas le sens de : je l'ai déjà demandé.

16:54:45:Q= je vous ai déjà demandé de faire la recherche.

16:54:50:R=Je ne comprends pas le sens de : je vous ai déjà demandé de faire la recherche.

16:55:24:Q= je le veux.

16:55:29:R=Désolé, je n'arrive pas à saisir votre propos.

16:55:54:Q= je veux effectuer une recherche selon ces critères.

(...)

Capture 7 : Un exemple d'échec dialogique avec le système Halpin-Recueil

Pour améliorer cette interaction entre l'homme et la machine, il faut prendre en compte les erreurs et les échecs par rapport au dialogue idéal [NERZIC 93]. Taleb explique : "*Dans un DHM, les écarts sont inévitables et ils se produisent pour diverses raisons. Sachant qu'on ne peut pas réduire leur occurrence, il faut concevoir des machines capables de les gérer. Ainsi, les systèmes doivent mémoriser les écarts et leurs stratégies de récupération, afin de pouvoir les réutiliser lors de l'apparition de cas similaires. Le dialogue est un processus d'apprentissage, les expériences doivent profiter aux expériences qui suivent.*" [TALEB 96].

5.4.1.2 Les tours de parole

Le corpus Halpin-Recueil compte 980 fichiers de DHM. Le Tableau 3 présente des statistiques relatives au nombre de questions et de réponses contenues dans les fichiers du corpus. La différence entre le nombre de questions et de réponses provient essentiellement du fait que certaines réponses sont actionnelles plutôt qu'informationnelles¹⁶.

	NOMBRE DE QUESTIONS	NOMBRE DE RÉPONSES
Moyenne	6,65	5,56
Écart type	5,51	4,97
Maximum	63	54
Minimum	1	1

Tableau 3 : Statistiques à propos du nombre de questions (homme) et de réponses (machine) par dialogue

¹⁵ (...) signifie que nous avons coupé une partie du dialogue

¹⁶ Typiquement, c'est le cas où la machine demande " Voulez-vous effectuer cette requête ? " ; l'utilisateur répond " oui " et en réponse à cette acceptation, la machine présente à l'écran les informations demandées. Il est alors possible de cliquer sur les hyperliens et d'obtenir d'autres informations. Ces données n'apparaissent pas comme une réponse dans le fichier de dialogue.

Nous ne pouvons pas comparer ces résultats statistiques avec d'autres travaux de DHM. D'une part, parce que chaque type de tâche entraîne des dialogues différents (et ici la tâche était particulièrement inhabituelle), et d'autre part, parce que les capacités et les attitudes des interlocuteurs (de la machine notamment), vont amener des dialogues totalement différents d'un système à l'autre.

Il est tout aussi difficile de comparer nos résultats avec un corpus de dialogues Homme-Homme. Là encore, principalement parce que les tâches sont différentes. Colineau donne par exemple, dans sa thèse, des nombres d'énoncés beaucoup plus élevés que les nôtres : « *Le nombre d'énoncés par dialogue varie de 46 énoncés à 232 énoncés avec une moyenne de 80 à 90 énoncés par dialogue (pour le corpus Observation) ; et (...) les dialogues analysés sont en moyenne plus courts, on observe des dialogues qui varient de 26 à 128 énoncés (pour le corpus Test) .* » [COLINEAU 97].

En bref, compte tenu de la simplicité de la tâche à accomplir (il ne fallait pas rebuter les usagers qui se connectaient sur le Web pour participer à cette expérience), cette moyenne de 11 énoncés (6 pour l'homme et 5 pour la machine) nous paraît normale pour une interaction hyperdialoguée, c'est-à-dire dans laquelle l'utilisateur dialogue avec la machine mais aussi clique sur des hyperliens pour obtenir des informations. La Figure 18 ci-après présente le nombre de tours de parole comptabilisés sur les 200 premiers fichiers du corpus Halpin-Recueil. On observe des pics qui correspondent à des fichiers contenant un grand nombre d'échanges dans les dialogues.

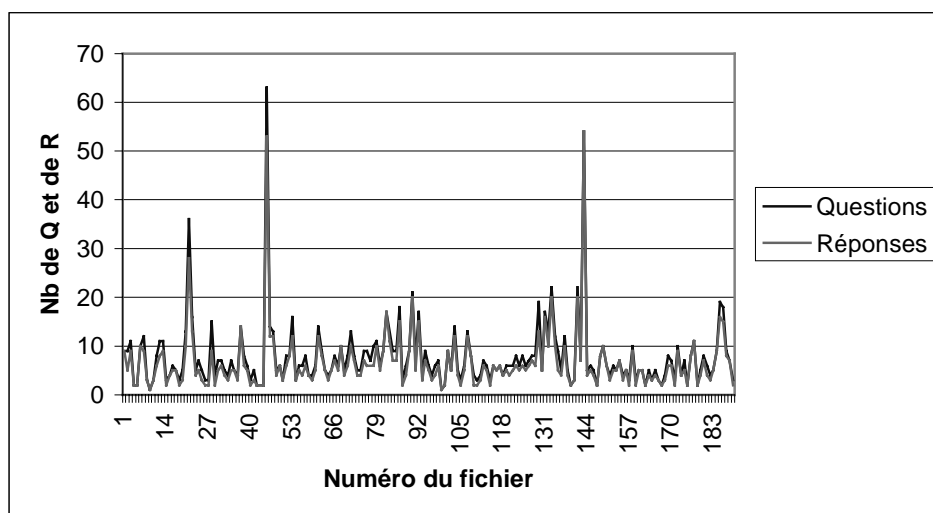


Figure 18 : Tours de parole dans le corpus Halpin-Recueil

Le nombre maximum de questions (c'est-à-dire d'énoncés de l'utilisateur) pour un seul dialogue est de 63 interventions¹⁷, alors que la moyenne est d'environ 7 questions par fichier. Le maximum parmi les réponses (énoncés de la machine) est de 54, la moyenne étant autour de 6, par fichier.

5.4.1.3 Les fréquences lexicales

En linguistique, l'expression « mot vide » désigne les mots qui ne sont pas des mots « de contenu » sémantique. En français, ce sont les articles, les prépositions, etc. S'il est vrai qu'ils ne sont pas réellement indispensables au traitement automatique global de la langue lorsqu'ils sont pris isolément, des travaux ont cependant montré qu'ils apportent des informations sémantiques non négligeables en contexte [VEILEX 85].

Par opposition aux mots vides, on appelle "mots pleins" les mots porteurs d'information (noms communs, verbes, etc.). En plaçant en abscisses, les mots d'un corpus rangés dans leur ordre décroissant d'apparition, et en ordonnées les fréquences d'apparition de ces mots, on obtient une

¹⁷ En fait, c'est un utilisateur du LIMSI (Laboratoire d'Informatique pour la Mécanique et les Sciences de l'Ingénieur) à Orsay, qui s'est connecté pendant près d'une heure et qui a testé la robustesse du système. Voici des extraits de leur dialogue :

m79.limsi.fr LE Mon Mar 2 18:31:41 MET 1998
18:31:55:Q= bonjour
18:31:59:R=Ah, enfin un humain qui dit "Bonjour" à une machine.
18:31:59:R=Dialoguons ensemble encore un peu...
18:32:34:Q= mais certainement, avec plaisir.(...)
18:42:18:R=Les critères que vous avez déjà trouvés sont :
18:42:18:R=Le sujet
18:42:18:R=La date de publication
18:42:18:R=Les mots-clefs de ce livre sont : Victime; Victimologie; Victimation; Indemnisation; Peine; Aide aux victimes; Entourage de la victime; Préjudice; Médiation
18:42:18:R=La date de publication est : 1991.
18:43:04:Q= 1672 réponses !
18:43:10:R=Désolé, je n'arrive pas à saisir votre propos.
18:43:44:Q=normal, c'est moi qui saisi mon propos au clavier.(...)
18:43:49:R=Voulez-vous effectuer une recherche selon ce(s) critère(s)?
18:44:06:Q= et si je dis non ?
18:44:06:R=D'accord, on ne fait pas de recherche pour l'instant.
18:44:11:R=Dialoguons ensemble encore un peu...
18:44:40:Q= quel temps fait-il à grenoble ce soir ?
18:44:44:R=Désolé, je n'arrive pas à saisir votre propos.
18:45:47:Q= décidément, vous voulez dialoguer mais vous l'avez pas beaucoup de conversation (à part les mots-clefs compulsifs...)
18:45:53:R=Les mots-clefs de ce livre sont : Victime; Victimologie; Victimation; Indemnisation; Peine; Aide aux victimes; Entourage de la victime; Préjudice; Médiation
18:45:53:R=Voulez-vous effectuer une recherche selon ce(s) critère(s)?
18:46:23:Q= les voila ! ne vous l'avais-je point prédit ?
18:46:32:R=Décidément, je ne comprends toujours pas ce que vous me dites.
18:47:59:Q= bon et bien allons droit au but : quel est l'auteur du livre sur la victimologie écrit en 1991 dont je dois trouver l'isbn ? est-ce assez simple ?
18:48:04:R=L'Auteur de cet ouvrage est : KAISER, Günther. (...)

courbe, connue sous le nom de « courbe de Zipf¹⁸ » [RAMAUX 95]. George Kingsley Zipf a étudié la distribution des fréquences d'apparition des mots d'une langue. Il observe que les corpus comportent généralement deux classes de mots : des mots apparaissant extrêmement souvent et d'autres, beaucoup moins utilisés.

Ainsi, dans les 298 000 mots du *Ulysse* de James Joyce, Zipf a découvert que :

- le 10e mot le plus employé est utilisé 2653 fois
- le 100e mot le plus employé est utilisé 265 fois
- le 1000e mot le plus employé est utilisé 26 fois

Dans chaque cas, le produit du rang par la fréquence égale plus ou moins 26500, et l'on constate que l'équation Rang x Fréquence = Constante.

La Figure 19 ci-dessous présente en abscisses le rang des 100 premiers mots rencontrés dans le corpus Halpin-Recueil, et en ordonnées, la fréquence d'apparition de ces mots. Cette courbe suit une loi de Zipf, comme beaucoup de corpus de base étudiés dans notre équipe (notamment BREF – réalisé au LIMSI-CNRS, ce corpus contient des articles lus, tirés du quotidien "Le Monde" – [GAUVAIN et al. 90], [LAMEL et al. 91]). Cependant, les premiers substantifs rencontrés sont tous relatifs au domaine (recherche documentaire), ce qui confirme l'hypothèse d'une forte influence de la tâche sur le vocabulaire employé par les sujets.

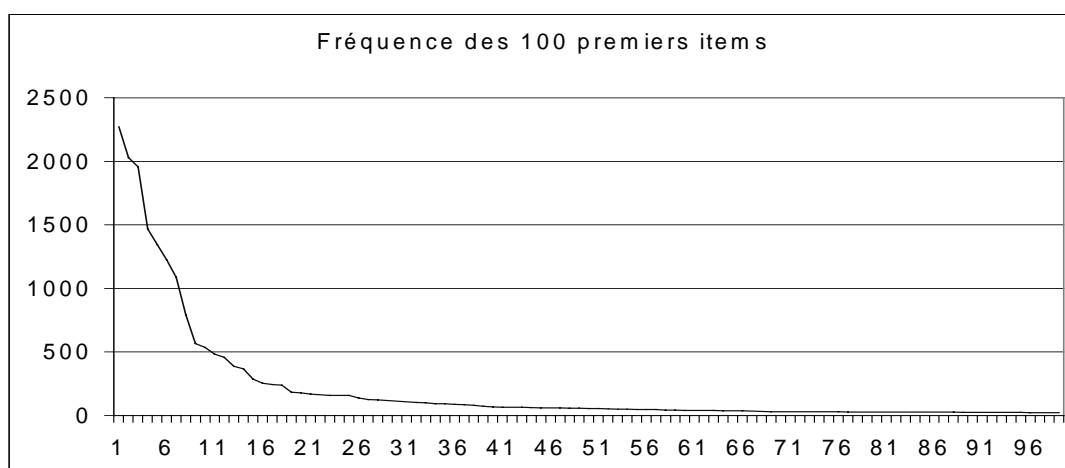


Figure 19 : Fréquence d'apparition des 100 premières occurrences du corpus Halpin-Recueil

¹⁸ Une courbe de Zipf exprime soit a) une relation entre la fréquence d'un événement et le nombre d'événements différents apparaissant avec cette fréquence, soit b) une relation entre la fréquence d'occurrence d'un événement et le rang qu'il occupe quand les événements sont ordonnés en fonction de la fréquence d'occurrence.

A propos de listes orthographiques, Catach présente des résultats quantitatifs pour les mots et leurs formes fléchies les plus fréquentes en langue française [CATACH 85]. Son classement par ordre de fréquence décroissante montre une forte apparition d'articles et de l'auxiliaire « être ». Des résultats similaires ont été obtenus avec le corpus Halpin. Le Tableau 4 donne les 25 premières occurrences avec leur rang et leur fréquence d'apparition.

Rang	Lemme	Fréquence	Rang	Lemme	Fréquence
1	le	2273	14	isbn	369
2	?	2031	15	parler	287
3	être	1958	16	je	254
4	livre	1468	17	avoir	246
5	de	1345	18	quoi	238
6	quel	1226	19	en	184
7	ce	1089	20	qui	180
8	oui	793	21	que	168
9	de=le	566	22	sujet	163
10	auteur	537	23	ouvrage	160
11	un	483	24	numéro	159
12	il	458	25	et	159
13	titre	389			

Tableau 4 : Fréquence des 25 premières occurrences du corpus Halpin-Recueil

En analysant un corpus de DHM, Fréchet publie des résultats proches de travaux de référence du domaine [GOUGENHEIM et al. 56], malgré le caractère finalisé du vocabulaire employé par les sujets. Elle dresse également une taxonomie des verbes en fonction des catégories sémantiques de l'application [FRECHET 88]. Nous nous en sommes inspiré pour la déclaration des concepts à identifier dans chaque énoncé des usagers.

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 ? » dans lesquelles le verbe peut fournir une information sémantique précieuse (cf. Figure 20 ci-après).

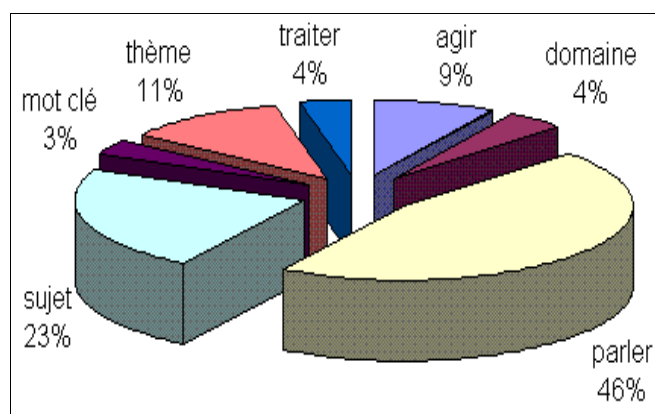


Figure 20 : Le concept du « thème »

5.4.2 Observations similaires aux corpus de DHM classiques

Après cette première présentation du corpus Halpin-recueil, nous allons montrer dans cette section que l'on retrouve dans ces dialogues des phénomènes déjà observés avec d'autres systèmes de DHM, notamment la co-construction d'un vocabulaire commun, et l'influence de la machine sur le comportement langagier de l'homme. Il apparaît dans ce corpus, comme dans d'autres travaux du domaine, que le dialogue entre l'homme et la machine, pour une tâche finalisée, ne peut pas se simplifier à des enchaînements de questions et de réponses.

5.4.2.1 Le dialogue : plus qu'un échange de questions et de réponses

Un système de DHM doit être différencié des applications de communications proches d'un DHM, dans lesquelles l'utilisateur n'est pas maître de ses actions. La Figure 21 ci-après, présente le principe de fonctionnement d'une application utilisant des technologies vocales¹⁹. Dans cet exemple, la machine utilise la reconnaissance vocale pour identifier des mots-clés dans l'énoncé du client, recherche dans la base de données les informations demandées et synthétise une réponse orale.

En y regardant de plus près, on s'aperçoit qu'outre les aspects acoustiques, en entrée et en sortie du système, la réussite ou l'échec d'une requête dans ce genre d'application est fortement couplé à la puissance de compréhension de ce qui a été dit par l'utilisateur. Dans cet exemple de renseignement téléphonique (Figure 21), si les champs *rubrique* et *ville* sont facilement remplis avec les données « cinéma » et « Guingamp », le renseignement du champ *date* donne lieu à un calcul. En effet, le client n'a pas donné la date en valeur absolue, mais en valeur relative (« demain »).

Ce genre d'application ressemble à du DHM, mais l'utilisateur se contente de donner l'information demandée par la machine, sans jamais pouvoir dialoguer spontanément, comme il le ferait avec un interlocuteur humain. En effet, le LN est inévitablement ambigu. C'est cette ambiguïté même qui donne au LN sa portée et sa richesse conceptuelles, car chaque mot contient des sens multiples et, par conséquent, des concepts multiples. « *Le mot n'est pas simplement une idée ; c'est un petit univers d'idées. Bien plus, chaque phrase exprimée oralement contient la semence d'une discussion, non seulement à cause de son ambiguïté mais parce que tout ce qu'elle affirme implique dans le temps sa négation et son contraire.* » [POSTMAN 79].

¹⁹ Source : « France Télécom en actions, Le magazine du Club France Télécom », N°7, Juillet 1999.



Figure 21 : Principe de fonctionnement d'une application utilisant des technologies vocales

L'homme, par nature, a la faculté de *computer*²⁰, parfois sans même savoir quels sont les mécanismes en jeu dans ses raisonnements. Le Moigne explique « *Computer, ce n'est pas seulement calculer arithmétiquement des nombres, c'est, très généralement, manipuler et traiter des symboles.* » [LE MOIGNE 86].

Dans sa thèse, Chanet pose le problème de l'identification de la demande dans un dialogue finalisé. Elle dit que " *Tous les énoncés ne sont pas des demandes. Dans le domaine de la recherche de renseignements, le dialogue est conduit par l'utilisateur. Si cet utilisateur est a priori susceptible de formuler des demandes, il peut aussi apporter de l'information au système, ou l'utiliser à des fins autres qu'informatives.* " [CHANET 96].

²⁰ A prononcer à la française : c'est ici le verbe français, et non pas le substantif anglais.

Dans notre corpus, on retrouve cette propriété à plusieurs reprises. La Capture 8 est un exemple intéressant de ce genre d'apport d'information de la part de l'utilisateur. Avec l'énoncé "cherche moi le bon !", on voit qu'il voudrait faire faire une tâche à l'ordinateur, mais en fait, il n'apporte aucune information permettant de le faire. Plus loin, l'utilisateur aurait pu se contenter de donner l'ISBN sans autre commentaire, mais au lieu de cela, il préfère informer son interlocuteur de ses avancées (" aye, j'ai trouvé ! ").

cybele.imag.fr LE Thu Dec 11 09:42:03 MET 1997
(...)
09:44:04:Q=quelle est la date de sortie du livre
09:44:09:R=La date de publication est : 1993.
09:44:09:R=Voulez-vous effectuer une recherche selon ce(s) critère(s)?
09:44:32:Q=oui
09:48:40:Q=cherche moi le bon !
09:48:44:R=Je ne comprends pas le sens de : cherche moi le bon !
09:49:11:Q=quelle est l'isbn du livre que je recherche?
09:49:16:R=Ca, c'est ce que vous devez trouver !
(...)
09:51:52:Q=aye, j'ai trouvé ! 2840250535

Capture 8 : L'utilisateur n'exprime pas que des demandes

Les extraits de dialogues présentés dans la Capture 9, ci-après, montrent que certains usagers considèrent la machine comme un interlocuteur très coopératif et lui expliquent ce qu'ils attendent de manière explicite.

isis.imag.fr LE Thu Dec 11 09:48:38 MET 1997
09:49:53:Q=y a-t-il un rapport entre l'isbn d'un livre et ce dont il parle ?
(...)
09:52:25:Q=je veux que soient affichés les liens sur les livres qui traitent de la police et de la justice
(...)
09:56:13:Q=je veux compléter les critères avec un critère sur l'auteur
(...)
10:00:12:Q=je veux voir réaffichée la liste de tout à l'heure avec les critères croisés sur le sujet et sur l'auteur (...)
10:07:14:Q= retour à la liste précédente
(...)
10:11:06:Q= je veux effectuer une recherche selon le sujet, l'auteur, et la date de publication de ce livre (...)
10:18:06:Q= pourquoi aucun livre ne correspond-il à ces critères ?
(...)
10:19:45:Q= je voudrais reprendre la recherche sur le nom de l'auteur et le sujet
(...)
10:23:03:Q= qu'est-ce au²¹ détermine l'ordre des livres dans la liste ?

Capture 9 : L'utilisateur s'exprime avec la machine comme avec un interlocuteur humain

²¹ Citation sans correction de la faute de frappe "au" pour "qui".

5.4.2.2 Co-construction d'un vocabulaire commun : l'homme s'adapte à la machine et vice-versa.

Un dialogue ne se compose pas uniquement de questions et de réponses. Des énoncés font avancer le dialogue sur un axe régissant (le cours normal de la conversation), tandis que d'autres le font dévier sur un axe incident (des demandes de précision ou de reformulation par exemple) [LUZZATI 95]. Dans l'une comme dans l'autre de ces phases du dialogue, l'homme et la machine vont s'approprier des connaissances, grâce à l'apprentissage des informations données par l'autre. C'est ce que les chercheurs de l'école de Palo-Alto (Californie) désignent comme le langage commun des allocutaires : « *la communication avec l'autre passe avant tout par l'emploi d'un langage commun* » [WATZLAWICK 80].

dunia.imag.fr LE Wed Dec 10 19:02:40 MET 1997
19:10:45:Q=quelle est l'année ?
19:10:50:R=La date de publication est : 1993.
19:10:50:R=Voulez-vous effectuer une recherche selon ce(s) critère(s)?
19:11:02:Q=oui
19:12:08:Q=rechercher 1993 et droit
19:12:13:R=La date de publication est : 1993.
19:12:13:R=Voulez-vous effectuer une recherche selon ce(s) critère(s)?
19:12:55:Q=date droit editeur
19:13:00:R=La date de publication est : 1993.
19:13:00:R=L'Editeur est : Paris : L'Harmattan, 1993.
19:13:00:R=Voulez-vous effectuer une recherche selon ce(s) critère(s)?
(...)

Capture 10 : L'utilisateur utilise des informations données par la machine

L'extrait de dialogue présenté dans la Capture 10 montre que l'utilisateur utilise dans une de ses questions une information que la machine vient de lui fournir (Q=rechercher 1993 et droit). La machine comprend que 1993 est une date, et confirme la date de publication de l'ouvrage. Cette co-construction d'un vocabulaire commun permet un dialogue plus riche, du fait que l'homme prend conscience que la machine a interprété son énoncé convenablement.

Dans une approche similaire à la nôtre, Lehuen proposait également un système de DHM appliqué à la recherche documentaire, pour illustrer son modèle hypothético-expérimental [LEHUEN 97]. Son système, COALA, pose des hypothèses au fur et à mesure du dialogue, et les valide ou les infirme selon la suite des événements. Ainsi la machine apprend à chaque nouvelle interaction avec l'utilisateur.

Parfois, l'homme devance la machine dans son acquisition de renseignements. La Capture 11, ci-après, montre que c'est l'utilisateur qui a pris l'initiative dans ce dialogue en demandant à la machine " ce livre a-t-il pour titre "derniers domiciles connus..." ? ". L'ordinateur doit en déduire que l'utilisateur possède déjà cette information.

samy.inria.fr LE Mon Dec 15 10:42:07 MET 1997
 10:43:17:Q= quel est le theme general du livre?
 10:43:22:R=Les mots-clefs de ce livre sont: Logement; Habitat; Architecture; Evolution
 10:43:22:R=Voulez-vous effectuer une recherche selon ce(s) critère(s)?
 10:43:51:Q= oui
 10:45:23:Q= qui a ecrit ce livre?
 10:45:28:R=L'Auteur de cet ouvrage est : LEGER, Jean-Michel
 10:45:28:R=Voulez-vous effectuer une recherche selon ce(s) critère(s)?
 10:47:45:Q= ce livre a-t-il pour titre "derniers domiciles connus..."?
 10:47:49:R=Le titre de ce livre est : Derniers domiciles connus : Enquêtes sur les nouveaux logements 1970-1990 Jean-Michel Léger.

Capture 11 : L'utilisateur anticipe des informations données par la machine

Pour analyser le corpus Halpin, nous avons développé des programmes qui scrutent les fichiers de dialogues de manière automatique ; c'est-à-dire que les données statistiques sont mises à jour à chaque nouvel ajout de fichiers dans le répertoire concerné. Nous présentons ci-dessous quelques résultats qui nous paraissent pertinents à propos du DHM.

Les termes utilisés par les usagers pour essayer d'identifier les différents critères du jeu (auteur, titre, thème, éditeur, date de publication, collection, etc.), sont examinés ci-après.

Dans le Tableau 5, nous présentons quelques concepts associés aux critères de recherche. Dans chaque colonne de concepts se trouve le nombre de fois qu'un terme associé à ce concept a été détecté (chaque terme n'étant comptabilisé qu'une fois par fichier).

Concept AUTEUR	Concept THEME	Concept DATE	Concept EDITEUR	Concept LIVRE
auteur : 438 (48.83)	parler : 243 (27.09)	date : 73 (8.14)	éditeur : 100 (11.15)	livre : 698 (77.81)
nationalité : 13 (1.45)	sujet : 150 (16.72)	année : 66 (7.36)	édition : 44 (4.91)	ouvrage : 120 (13.38)
homme : 11 (1.23)	thème : 49 (5.46)	quand : 28 (3.12)	publication : 38 (4.24)	roman : 55 (6.13)
femme : 8 (0.89)	s'agir : 46 (5.13)	récent : 15 (1.67)	publier : 28 (3.12)	bouquin : 21 (2.34)
personne : 6 (0.67)	domaine : 26 (2.90)	siècle : 5 (0.56)	référence : 9 (1.00)	document : 10 (1.11)
	traiter : 17 (1.90)	époque : 4 (0.45)		
	thématique : 6 (0.67)	contemporain : 4 (0.45)		
	propos : 4 (0.45)	période : 2 (0.22)		
	résumé : 4 (0.45)	avant : 2 (0.22)		
	contenu : 4 (0.45)	jeune : 2 (0.22)		
	matière : 3 (0.33)	après : 1 (0.11)		
	sommaire : 1 (0.11)	vieux : 1 (0.11)		
	objet : 1 (0.11)			

Tableau 5 : Quelques concepts associés aux critères de recherche

Le taux entre parenthèse est relatif au nombre total de fichiers du corpus (896 pour cette étude). On voit, par exemple, qu'environ la moitié des fichiers du corpus comportent au moins un fois le terme « auteur » (48.83% du corpus). En valeur absolue, il a été employé 662 fois. Cela s'explique sans doute par le fait que les usagers commencent souvent par une requête à un seul terme, puis tentent ensuite de coupler des critères entre eux (auteur et date par exemple).

On retrouve dans 79,63 % des fichiers le concept « d'acceptation/confirmation ». Une très large partie de ces réponses sont des phrases où l'on peut déceler le mot « oui ». C'est en effet le cas pour 70,37 % des fichiers où l'on retrouve ce concept.

Les autres possibilités employées par les usagers pour ce concept sont variées : *ok, ouais, pourquoi pas, yes, d'accord, certainement, volontiers, avec plaisir*, sont quelques unes des variantes utilisées pour exprimer une approbation, comme le montre la Figure 22 ci-après.

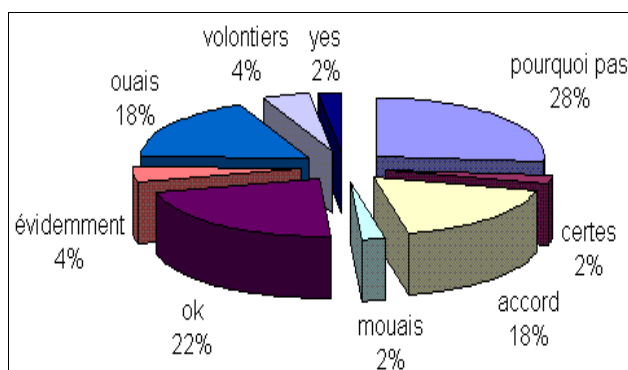


Figure 22 : Le concept « acceptation » (autre que oui)

Pour le concept de « négation », le vocabulaire se limite à trois termes : « non » (100 occurrences, soit 11,15 % de présence dans le corpus), puis « ne » et « pas » (49 et 56 occurrences, soit 5,46 et 6,24%). Dans les autres cas, les questions du système sont simplement ignorées. L'utilisateur parle alors d'autre chose et le système suppose qu'implicitement la réponse est négative.

5.4.2.3 Statistiques sur l'emploi de termes spécifiques à la tâche

Nous présentons maintenant des éléments statistiques permettant de déterminer comment le vocabulaire employé par la machine peut influencer celui de l'utilisateur. Certaines données ne sont pas présentées ici lorsque c'est la machine qui les emploie, parce qu'il peut s'agir de phrases non comprises qu'elle reformule. Par exemple, lorsque la machine répond " je ne comprends pas le sens de : est-ce que tu passes le test de Turing ? " (cf. annexe 4), le " tu " n'est pas employé à l'initiative de la machine, même si c'est bien elle qui produit cet énoncé.

TERME	Homme	%	Machine	%
ACCORD	10	1.11 %	106	11.82 %
ANNEE	66	7.36 %	4	0.45 %
AUTEUR	438	48.83 %	484	53.96 %
BOUQUIN	21	2.34 %	4	0.45 %
BRAVO			396	44.15 %
COMMENT	46	5.13 %	19	2.12 %
CRITERE	65	7.25 %	345	38.46 %
DATE	66	7.36 %	169	18.84 %
EDITEUR	100	11.15 %	20	2.23 %
LIVRE	698	77.81 %	856	95.43 %
OUVRAGE	120	13.38 %	625	69.68 %
POURQUOI	26	2.90 %	6	0.67 %
SCIENTIFIQUE	25	2.79 %	80	8.92 %
THEME	49	5.46 %		
TITRE	312	34.78 %	461	51.39 %
TU	45	5.02 %	14	1.56 %
VOUS	47	5.24 %	690	76.92 %

Tableau 6 : Quelques termes du vocabulaire employé par l'homme et par la machine

On note par exemple dans le Tableau 6 que le mot "auteur" est utilisé en proportion équivalente d'un côté comme de l'autre. La machine énonce plus souvent le mot "date" que l'homme car ce dernier peut formuler sa question de plusieurs façons, sans pour autant utiliser le mot date, tandis que la machine l'intègre forcément dans sa réponse. Pour repousser cette limite, on pourrait utiliser un thesaurus afin d'enrichir le vocabulaire de la machine, mais se pose alors le délicat problème du choix des termes à utiliser dans telle ou telle situation.

On observe également que le terme "bravo" apparaît dans environ 45% des fichiers du corpus. Cela signifie que cette part des usagers a accompli sa tâche correctement, c'est-à-dire qu'elle a retrouvé le numéro ISBN du livre sélectionné par la machine. En étudiant le corpus, on peut difficilement dire pourquoi l'autre moitié n'y est pas parvenue, mais on note quelques raisons possibles : arrêt de la tâche en cours de la part de l'utilisateur, lenteur du réseau Internet, manque de pertinence dans les réponses de la machine (vocabulaire trop restreint), etc.

Enfin, nous avons remarqué, comme cela a été observé dans d'autres corpus [MOREL 89], des phénomènes de mimétisme de l'utilisateur par rapport à la machine. Non pas en observation directe, puisque que les expériences se déroulaient de manière anonyme sur le Web, mais indirectement, au travers du vocabulaire employé par les utilisateurs. Par exemple, il est flagrant que le mot « critère » est pratiquement toujours employé par l'utilisateur, seulement après que la machine l'ait elle-même utilisé, dans une de ces réponses.

5.4.2.4 Les verbes utilisés par les utilisateurs

Un autre outil d'analyse a été développé afin d'étudier la composition du corpus en termes 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 7 montre la fréquence d'apparition des verbes les plus utilisés par les usagers. On note bien entendu une importante présence des auxiliaires *être* et *avoir* mais on retrouve également certains verbes de cognition comme *chercher*, *rechercher*, *connaître*, *trouver*, etc.

être	1365	traiter	23
parler	221	rechercher	21
avoir	177	aller	19
vouloir	73	dire	19
agir	37	chercher	16
pouvoir	37	titrer	14
faire	36	connaître	12
écrire	33	livrer	12
trouver	33	appeler	11
effectuer	30	poser	10
comprendre	25	devoir	8
donner	24		

Tableau 7 : Fréquence d'apparition des verbes les plus utilisés par les utilisateurs

A noter également, la présence d'éléments inattendus comme « livrer » ou « titrer ». Cela s'explique par une confusion (verbe/substantif) du lemmatiseur pour les termes « livre » et « titre ».

Nous avons vérifié sur le serveur de Xérox cette erreur, et effectivement, comme le montre le Tableau 8 ci-après, dans la phrase « je veux le titre », *titre* est bien analysé comme un nom, tandis que dans « donne-moi le titre », il est proposé comme un verbe, à la 3^{ème} personne du singulier.

je veux le titre			donne-moi le titre		
je	je	+PRON_P1P2	donne-	donner	+VERB_P1P2
veux	vouloir	+VERB_P1P2	-moi	moi	+PC
le	le	+DET_SG	le	le	+PC
titre	titre	+NOUN_SG	titre	titrer	+VERB_P3SG

Tableau 8 : Erreur de lemmatisation : titre vs titrer

5.4.2.5 Les substantifs utilisés par les utilisateurs

Le Tableau 9 dresse une liste non exhaustive des substantifs classés dans un ordre décroissant de fréquence d'apparition. On voit qu'ils sont majoritairement en rapport avec la tâche à accomplir proposée.

livre	1081	publication	44
oui	551	langue	42
auteur	411	éditeur	36
isbn	267	bibliothèque	33
titre	266	question	31
numéro	128	français	29
ouvrage	106	parution	29
recherche	83	roman	28
date	76	mot	27
critère	70	domaine	23
sujet	66	bouquin	22
année	55	sociologie	21
thème	51	chiffre	20
nom	49		

Tableau 9 : Fréquence d'apparition des noms les plus utilisés par les utilisateurs

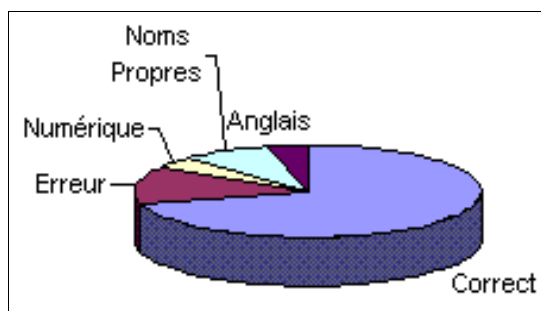


Figure 23 : Vocabulaire employé dans les dialogues Homme-Machine du corpus Halpin-Recueil

Le vocabulaire global (cf. Figure 23) utilisé par les usagers est composé de 527 termes différents : 70,78% sont des mots corrects de la langue française (on trouve par exemple « chat » dans l'énoncé « je donne ma langue au chat ») ; 13,09% sont des erreurs de frappe ou des mots qui n'existent pas ; 8,54% sont des noms propres et 3,80% des mots anglais. On retrouve également des expressions numériques dans le corpus Halpin (lorsque les utilisateurs donnent l'ISBN d'un ouvrage, le plus souvent).

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'explique Fréchet : « Les onze autres termes à fréquence élevée peuvent être des termes propres au corpus d'oral finalisé que nous étudions : présentatif, appuis du discours et connecteur pragmatique, des

termes propres à l'application (« fenêtre », « signal », « sur ») ou des termes ayant un rapport avec la tâche (« pouvoir », l'auxiliaire aller marquant l'intentionnalité du locuteur) » [FRECHET 88].

* * *

La plupart des résultats présentés ci-dessus sont donc relativement similaires aux observations effectuées avec d'autres systèmes de DHM. Nous exposons à présent des éléments qui nous semblent originaux, dans le comportement (notamment langagier) des utilisateurs, et que nous pensons être induits par le type de support utilisé pour cette communication homme-machine.

5.4.3 Observations différentes par rapport aux corpus de DHM classiques

Compte tenu du support utilisé, le temps nécessaire au recueil d'un corpus conséquent est très rapide grâce à cette technique, comparé à d'autres expériences réalisées en protocole du type Magicien d'Oz [ZANELLO 97], [CATINIS 98]. En quelques semaines, il est donc possible d'obtenir des centaines de dialogues exploitables pour étudier le DHM. Il est alors souhaitable de concevoir et de réaliser des programmes informatiques permettant d'exploiter les données obtenues.

L'analyse ci-après a pour but de rendre compte du vocabulaire employé et des stratégies développées par les utilisateurs qui se sont connectés au système. Grâce à un outil informatique que nous avons développé (en Perl) pour analyser les dialogues obtenus sur le Web, on observe des résultats quantitatifs : d'une part le nombre total d'occurrences pour un terme spécifié, et d'autre part le nombre *restreint* d'occurrences (c'est-à-dire qu'un terme n'est comptabilisé qu'une fois, même s'il apparaît plusieurs fois dans le même fichier).

Les résultats suivants ne font apparaître que le nombre *restreint* d'occurrences, car l'on s'intéresse aux termes qui reviennent souvent, sur l'ensemble du corpus, et non pas à des épiphénomènes : un terme pourrait être utilisé plusieurs fois par seulement quelques usagers, il n'en serait pas pour autant représentatif, au niveau du corpus global.

Les concepts d'ouverture et de clôture²² de dialogue (cf. Tableau 10) sont peu présents dans ce corpus. On observe, par exemple, seulement 16 « bonjour » dans les 898 fichiers de dialogues étudiés.

²² Parfois l'utilisateur prend congé de manière humoristique, comme le montrent ces 2 exemples :
13:22:21:R=Dialoguons ensemble encore un peu...

Ce phénomène est sans doute dû, d'une part, au fait que le dialogue est à l'initiative de l'utilisateur, et que ce dernier ne prend quelquefois pas le temps de saluer son interlocuteur (« quel intérêt de saluer une machine ? » pensent certains), préférant directement entrer dans le vif du sujet. D'autre part, une fois que l'ordinateur s'est présenté, les utilisateurs s'imaginent sans doute qu'il est en attente d'une action à effectuer, et entrent directement dans le cœur de la conversation relative à la tâche. 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.

Concept « Ouverture de dialogue »	Concept « Clôture de dialogue »
bonjour : 16 (1.78 %)	bye : 19 (2.12 %)
bonsoir : 3 (0.33 %)	stop : 2 (0.22 %)
salut : 14 (1.56 %)	finir : 1 (0.11 %)
	fin : 1 (0.11 %)
	arrêter : 6 (0.67 %)
	partir : 6 (0.67 %)
	terminer : 1 (0.11%)
	tchao : 1 (0.11%)
	abandonner : 5 (0.56 %)

Tableau 10 : Les concepts d'ouverture et de clôture de dialogue

Cependant, on trouve quelques dialogues dans lesquels une formule d'ouverture de dialogue est donnée par l'utilisateur. Elle peut être indépendante, ou bien liée à la première question que l'utilisateur souhaite poser à la machine, comme le montrent les extraits de la Capture 12 ci-dessous.

<p><i>m79.limsi.fr LE Mon Mar 2 18:31:41 MET 1998</i> 18:31:55:Q= bonjour 18:31:59:R=Ah, enfin un humain qui dit "Bonjour" à une machine. 18:31:59:R=Dialoguons ensemble encore un peu... 18:32:34:Q= mais certainement, avec plaisir.</p> <p><i>orpee.imag.fr LE Thu Jan 15 13:33:30 MET 1998</i> 13:34:17:Q= salut mec 13:34:21:R=Ah, enfin un humain qui dit "Bonjour" à une machine. 13:34:21:R=Dialoguons ensemble encore un peu... 13:34:35:Q=volontiers, qui est l'auteur alors ?</p> <p><i>baba.cnrs-mrs.fr LE Fri Jan 16 18:11:22 MET 1998</i> 18:11:41:Q= bonjour ! 18:11:46:R=Ah, enfin un humain qui dit "Bonjour" à une machine. 18:11:46:R=Dialoguons ensemble encore un peu... 18:12:15:Q= a propos de quoi ?</p>
--

13:22:59:Q= c'est ça, c'est ça. demain. tu m'excuses, j'ai du lait sur le feu. atchao !

13:18:57:Q= lemmatisation, c'est une sorte de dératisation, c'est ça ?

13:19:03:R=Décidément, je ne comprends toujours pas ce que vous me dites.

13:22:16:Q= bon, ben moi je vais à la pêche. allez, salut, stucru !

leucippe LE Thu Feb 12 16:49:01 MET 1998

16:49:13:Q= bonjour mon gars

16:49:18:R=Ah, enfin un humain qui dit "Bonjour" à une machine.

16:49:18:R=Dialoguons ensemble encore un peu...

16:49:49:Q= que veux tu savoir ?

16:49:54:R=Je ne comprends pas le sens de : que veux tu savoir ?

16:50:08:Q=ah, ca commence mal

16:50:08:R=Tout commence par quelques chose, mais là je ne peux pas vous aider ...

16:50:13:R=Dialoguons ensemble encore un peu...

16:50:31:Q=bon alors, qui est l'auteur

* * *

sun2.lri.fr LE Mon Mar 23 10:30:45 MET 1998

10:31:29:Q= bonjour, quel est le format de l'ouvrage ?

macnkaoua.u-bordeaux2.fr LE Wed Jan 14 15:54:15 MET 1998

15:54:50:Q=bonjour, je voudrai connaître le numero isbn du livre en question ?

zig.inria.fr LE Tue Dec 16 16:26:17 MET 1997

16:26:39:Q=bonjour quel est ton nom ?

Capture 12 : Exemples d'ouverture de dialogue dans le corpus Halpin-Recueil

5.4.3.1 La mise en doute des informations données par la machine

En règle générale, l'homme est assez tolérant dans ses rapports avec la machine, mais parfois, comme dans l'extrait de dialogue de la Capture 13, il met en doute les capacités de la machine ou, tout au moins, son intention de coopérer.

Des énoncés comme « tu mens », « vérifie ta base de données », « vous êtes sûr ? » ont été observés dans notre corpus. Cela montre bien que l'homme dialogue avec le système de la même manière qu'il le ferait avec un humain ; en mettant parfois en doute les informations données, en demandant de les vérifier, et en essayant de jauger les limites du système.

aruba.inrialpes.fr LE Thu Nov 13 15:25:20 MET 1997

15:25:36:Q= quel est le titre de l'ouvrage ?

15:25:41:R=Le titre de ce livre est : Alternatives à la prison : Diffusion ou décroissance du contrôle social [Entrevue avec] Michel Foucault. Criminologie, Montréal : Presses Universitaires de Montréal, 1993, vol. 26, n°1, pp. 13-34

15:25:41:R=Voulez-vous effectuer une recherche selon ce(s) critère(s)?

15:26:42:Q= oui

15:28:04:Q= l'isbn de ce livre est 2760624420.

15:28:04:R=Désolé, ce n'est pas le bon ISBN.

15:28:09:R=Dialoguons ensemble encore un peu...

15:28:26:Q=non, tu mens.

(...)

Capture 13 : L'utilisateur doute des informations données par la machine

Dans ce cas précis, après vérification, il s'avère que c'est effectivement l'utilisateur qui avait raison. Il a donné le bon ISBN, mais pour une raison indéterminée, la machine « s'est trompée », et n'a pas accepté la (pourtant bonne) réponse.

Même si cette anecdote est sans doute due à une erreur de programmation de notre part, il n'en demeure pas moins intéressant de voir qu'ici, l'usager ne considère pas la machine comme un interlocuteur en qui l'on peut avoir une entière confiance. Notons cependant, comme nous l'indique la première ligne du dialogue, que cet usager s'est connecté depuis l'INRIA. C'est donc sans doute une personne ayant de bonnes connaissances en informatique. Un novice n'aurait peut-être pas osé remettre en cause la décision de la machine.

Même si ces phénomènes ont déjà été observés dans d'autres études, nous pensons que c'est une particularité de ce types de corpus de DHM, dans lesquels les utilisateurs sont distants (connexion sur le Web), non identifiables (un numéro de machine peut être détecté, mais la personne demeure anonyme), et non « surveillés » par un expérimentateur, que de comporter des éléments de contestation plus marqués, voire des insultes ou des plaisanteries, que dans d'autres types de corpus recueillis dans notre communauté scientifique.

5.4.3.2 La personnalisation

Les études de [CARBONELL et al. 83] ont montré que nous employons volontiers des références pronominales en LN. Ces procédés anaphoriques²³, tout comme les ellipses²⁴ ou les tropes²⁵ assurent la concision et la souplesse d'une LN. Il est donc intéressant, pour vérifier la validité de notre système de vérifier que ce corpus de DHM en langue naturelle recèle de tels éléments. A ce propos, Jacquet écrit « *On peut même penser que la perception d'autrui qui fonde l'une des régulations essentielles du dialogue est irréductible à la « perception » dont une machine est capable. (...) Cela dit, cette machine (faut-il le rappeler ?) est différente d'un autrui. (...) Peut-être que l'une des conditions indispensables au bon fonctionnement de l'illusion est le transfert des pronoms personnels ?* » [JACQUET 97].

²³ L'anaphore est un mot renvoyant à un membre de phrase déjà énoncé.

²⁴ L'ellipse est l'omission d'un ou de plusieurs mots d'un énoncé sans altération du sens de l'énoncé.

²⁵ Un trope est une figure rhétorique consistant à employer un mot ou une expression en les détournant de leur sens propre.

Le Tableau 11 donne la liste de quelques fréquences de pronoms du corpus, ainsi que leur rang par rapport à l'ensemble des occurrences rencontrées.

Pronom	Fréquence	Rang
Je	254	16
Tu	66	43
Il	458	12
Nous	8	181
Vous	63	44
On	25	94
Moi	58	49
Toi	7	222
lui	1	999
celui	3	471
ceci	1	1118
cela	2	650
leur	3	462
qui	180	20

Tableau 11 : Fréquence et rang des pronoms du corpus Halpin-Recueil

Jacquet dit que l'utilisation du « tu » signifie souvent que les sujets s'adressent clairement au compère humain, en situation de Magicien d'Oz. C'est-à-dire que le vouvoiement est remplacé par le tutoiement en signe de « reconnaissance », de coopération entre hommes. Dans notre expérience, la machine vouvoie l'utilisateur. Or, on retrouve dans les résultats autant de « tu » que de « vous ». Donc, soit certains sujets ont cru avoir affaire à un interlocuteur humain, soit, et c'est plus vraisemblable, ils dialoguent avec la machine comme avec un vrai partenaire.

Comme Chanet le souligne, un taux élevé d'impératifs et de « tu » dans un corpus peut s'expliquer si les utilisateurs sont des informaticiens : ils ont une grande habitude des langages de commande – lesquels comportent des impératifs – et aussi une relation assez familière à leur machine [CHANET 96]. Notons également la présence de « on » et de « nous », ce qui dénote qu'il puisse s'instaurer un rapport collaboratif entre les interlocuteurs.

Contrairement aux résultats d'expériences en Magicien d'Oz, où le « je » est marginal, ici, sa fréquence d'apparition est importante. On peut alors penser que l'utilisateur s'implique plus dans le dialogue. Par ailleurs, des phrases comme « je voudrais que tu ... », « je souhaite obtenir ... » sont courantes dans le corpus Halpin. Cela semble indiquer que l'être humain cherche à faire faire des actions à la machine.

Notons la forte apparition du pronom « il » dans ce corpus. En effet, ce pronom qui fait souvent référence à la « non personne » (comme dans « il fait beau »), arrive ici au 12^{ème} rang des mots du

corpus et avec la plus forte fréquence d'apparition parmi les pronoms (458 fois). En observant de plus près les énoncés des usagers, on note que le « il » apparaît majoritairement dans des phrases interrogatives (Cf. Capture 14) et quelques fois en élément anaphorique, se rapportant le plus souvent aux concepts « livre » et « auteur ».

11:36:29:Q=combien y a-t-**il** de livres en bibliothèque ?
 11:41:05:Q=l'auteur est-**il** un homme?
 01:00:55:Q=le livre est-**il** cher
 16:02:53:Q=a quelle époque le livre s'intéresse-t-**il** ?
 13:19:17:Q=l'isbn du livre est-**il** pair ?
 04:17:54:Q=le livre est-**il** en français
 15:05:00:Q=est-ce qu'**il** s'agit d'un roman ou d'un ouvrage à caractère technique?
 13:00:22:Q=ce livre contient-**il** des images ?
 13:05:29:Q=le titre est-**il** un oiseau sur la branche
 16:06:56:Q=dans quel pays le livre a-t-**il** été édité ?
 16:37:50:Q=peut-on disposer du nom de l'auteur de cet ouvrage ? si oui, quel est-**il** ?
 00:20:44:Q=de quoi traite-t-**il** ?
 11:17:48:Q=de quoi parle-t-**il** ?
 19:41:48:Q=quand a-t-**il** été écrit ?

Capture 14 : Quelques énoncés des usagers comportant le mot "il", (corpus Halpin-Recueil)

❖ **Résultats après plusieurs itérations du système Halpin-Recueil**

Le système de DHM Halpin se veut donc itératif et évolutif. En partant d'un état initial D0, le système est amélioré au fur et à mesure de l'analyse des résultats obtenus en utilisation réelle. Ainsi, il est possible d'observer les données du corpus au rythme des vagues d'expériences lancées sur le Web, et de comparer les résultats aux niveaux D1, D2, Dn, relativement les uns aux autres.

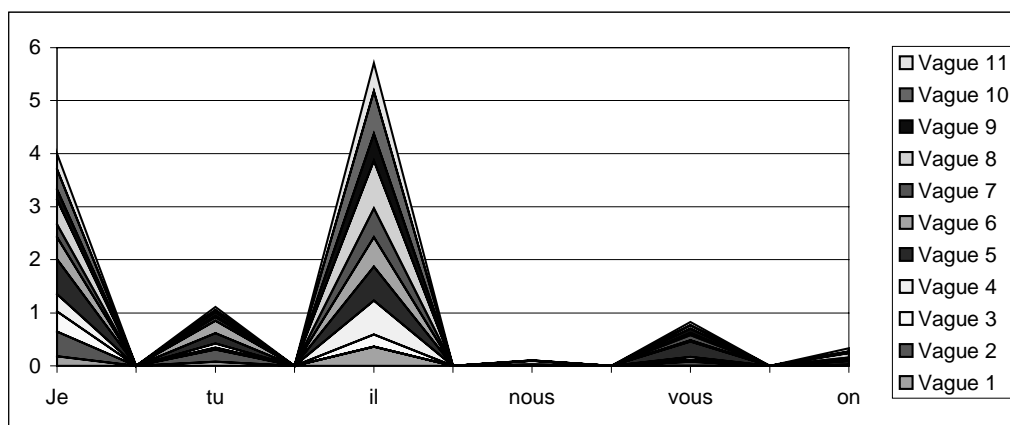


Figure 24 : Fréquences relatives de pronoms en fonction des différentes vagues d'expériences

La Figure 24 ci-dessus représente les fréquences relatives d'apparitions de pronoms sur les 11 premières vagues d'expériences. Cette fréquence relative pour une vague donnée est calculée en

rapportant la fréquence d'apparition du pronom étudié sur le nombre de fichiers pour ce sous-corpus. Globalement, on retrouve donc les mêmes résultats que dans le Tableau 11. On note par ailleurs que les données recueillies au cours des différents états (D0, D1, ..., D11) sont relativement homogènes, sauf pour le pronom « vous » dans la vague 5, avec une fréquence relative de 0.295, tandis que la moyenne des autres vagues pour ce pronom est de 0.052. Nous n'avons pas trouvé d'explications à ce phénomène.

5.4.3.3 Influence du support (Internet) sur le dialogue

Certains phénomènes ou comportements langagiers rencontrés dans nos corpus nous semblent être typiquement imputables au type de support utilisé dans cette communication. Le fait d'être connecté au réseau Internet influence, consciemment ou non, les comportements des usagers qui dialoguent avec un ordinateur distant.

Tout d'abord, il est très fréquent de voir apparaître des mots mal accentués, comme par exemple dans la phrase « 20:03:38:Q=qu'y a-t-il marqué sur la page de garde ? ». Nous pensons que cela n'est pas une erreur de frappe, mais un oubli volontaire de la part de l'utilisateur. En effet, nous savons que pour des raisons d'incompatibilités entre certains types d'ordinateurs (PC vs Mac, par exemple), les caractères accentués sont quelquefois mal retranscrits d'une machine à l'autre (le caractère diacritique est alors remplacé par un code). Une partie des utilisateurs du courrier électronique (E-mail) et de certains sites de « chat » (conversation/bavardage) ont pris volontairement l'habitude d'omettre l'accentuation dans leurs communications via Internet.

Ce phénomène semble donc ici se retrouver, dans les dialogues homme-machine que nous avons recueillis. Si pour l'homme, il est facile de corriger, sans même parfois s'en rendre compte, de tels défauts d'accentuation, il n'en va pas de même pour une machine, et cela risque de compliquer sa tâche de compréhension de l'énoncé fourni par l'utilisateur.

D'autres phénomènes de ce genre sont observables dans notre corpus. On y trouve, par exemple la contraction de mots et l'apparition de smileys.

<p><i>mendeleiev.cybercable.fr SUR Mozilla/4.0 (compatible; MSIE 4.01; Windows 95) LE Tue Oct 26 12:23:06 MET 1999</i> 12:23:07:M0=Que souhaitez-vous faire ? 1 : Retrouver un document que vous connaissez, 2 : Rechercher un ensemble de documents, ou 3 : Découvrir ce site ? 12:23:07:H0=2</p>
--

12:23:20:M1=Rechercher un ensemble de documents. D'accord. Souhaitez-vous que je vous aide à formuler votre requête ?
 12:23:27:H1=ok
 12:23:37:M2=Pouvez-vous me donner le thème recherché, entre guillemets s'il vous plaît ?
 12:23:59:H2=cherche "java"
 12:24:16:M3=J'ai trouvé 81 documents ayant comme thème java . Voulez-vous affiner votre requête ?
 12:24:50:H3=sur koi j'peux affiner
 12:25:06:M4=Afin d'affiner, vous devez choisir parmi les critères : Auteur, Titre, Thème, Année, ou Type de document.
 12:25:14:H4=theme
 12:25:40:M5=Le thème actuel est java . On pourrait y ajouter par exemple : réutilisation logiciel ; système réparti ; développement logiciel ; enseignement informatique ; ada ; . Pour ajouter un thème, donnez le entre guillemets, sinon répondez Non pour passer à autre chose.
 12:26:38:H5=tu es spécialiste java on dirait :-)

Capture 15 : Utilisation de mots contractés et de smileys dans un DHM sur Internet

La Capture 15 ci-dessus montre comment un internaute contracte le mot « quoi » dans la phrase « H3=sur koi j'peux affiner ». Il est assez facile pour un humain de comprendre ce genre de contraction, même si on les rencontre pour la première fois, parce que ces abréviations sont principalement d'ordre phonétique.

Ainsi, dans des discussions entre connectés sur Internet, on peut rencontrer par exemple :

- « T sur ? » (pour « t'es sûr ? »),
- « C koi ton pb ? » (pour « c'est quoi ton problème ? »),
- « Alors, koi 2 9 ? » (pour « alors, quoi de neuf ? »),
- « A+ » (pour « à plus tard »), etc.

En anglais, les usagers utilisent souvent des acronymes pour éviter d'avoir à taper toute une expression, comme le montre le Tableau 12 ci-dessous.

Abréviation	Correspondance	Traduction française
IMHO	In My Humble Opinion	A mon humble avis
J/K	Just Kidding	Je plaisante
AFK	Away from keyboard	Loin du clavier
BAK	Back at keyboard	De retour au clavier
BBL	Be Back Later	Je reviens plus tard
CUL8er	See you Later	A bientôt
OIC	Oh I See	Oh, je vois...
BTW	By The Way...	Au fait...
GTRM	Going To Read Mail	Je vais lire mes messages

Tableau 12 : Quelques abréviations anglaises courantes pour les dialogues sur Internet

C'est à la fois un moyen de gagner du temps, et une façon de revendiquer une appartenance à un groupe d'individus (seuls les anciens et les initiés comprendront ces messages).

De même, les smileys (« émoticônes » en français), sont des signes permettant de faire passer une humeur, un sentiment ou une émotion de manière écrite, sur Internet. C'est souvent un moyen de

remplacer le ton de la voix. Il faut pencher la tête sur le côté gauche pour comprendre la signification du smiley. Quelques smileys sont donnés à titre d'exemple dans le Tableau 13.

: -)	Le smiley le plus connu, souriant, utilisé pour indiquer un passage humoristique.
;-)	Le smiley clin d'œil, pour les remarques ironiques, par exemple.
:-(Le smiley triste : on n'a pas aimé ce qui vient de se dire, ou l'on est pessimiste, ou mélancolique.
: - o	Le smiley surpris, ou choqué.

Tableau 13 : Les smileys les plus utilisés sur Internet

Ainsi, dans le dialogue présentée en Capture 15, l'utilisateur dit à la machine « H5=tu es spécialiste java on dirait :-) ». Le smiley à la fin de la phrase est donc ici présent pour signifier une plaisanterie. Notons également que cet utilisateur n'a accentué ni le mot « thème » ni le mot « spécialiste ».

Notre système n'étant pas programmé pour interpréter les smileys et les contractions de mots, il ne pouvait pas tenir compte de ces éléments dans son interprétation.

Il serait intéressant, dans le cadre d'un dialogue sur le Web, comme le nôtre, qu'une machine puisse interpréter ces formes de communications, et en générer elle-même, si nécessaire.

Il semble donc, au vu de ces résultats, que certaines règles²⁶ ou coutumes applicables sur Internet, lorsque l'on dialogue entre hommes, se retrouvent dans une situation de dialogue avec une machine.

Enfin, nous pensons, mais cela reste à vérifier dans des travaux plus approfondis, que la longueur des phrases utilisées dans un DHM sur le Web, permet d'identifier un type d'utilisateur particulier : les novices auront tendance à faire de longues phrases, en donnant un grand nombre de détails à la machine, tandis que les experts se limiteront aux mots-clés qu'ils savent compris et bien interprétés par l'ordinateur.

Dans des cas particuliers, un expert fera une longue phrase, pour réduire le nombre d'interactions avec la machine, s'il sait qu'*a priori*, cette dernière est capable d'interpréter une série de mots-clés ou de concepts, sans avoir à interroger plusieurs fois l'utilisateur.

²⁶ Par exemple, la règle qui conseille d'éviter d'écrire un message tout en majuscule, car CELA DONNE L'IMPRESSION QUE L'ON CRIE.

En résumé :

Le système Halpin-Recueil a permis de recueillir un corpus de dialogues réels entre des utilisateurs et une machine, en utilisant le Web comme terrain d'expérimentation. Nous avons analysé environ un millier de fichiers dans lesquels des personnes volontaires dialoguent avec le système Halpin pour retrouver un document qu'il a tiré au hasard.

Grâce à ces expériences nous avons pu étudier des paramètres importants du DHM, comme le type de langage utilisé, le vocabulaire employé, l'enchaînement des tours de parole, l'importance d'une restitution sonore des réponses de la machine, etc.

Le point essentiel à retenir nous semble être l'importance de l'analyse sémantique pour la compréhension des énoncés de l'utilisateur. Nous avons vu, en effet, que les usagers utilisent pleinement la richesse de la langue naturelle pour dialoguer avec la machine.

Il nous paraît important de doter la machine, d'une part, d'un processus de reconnaissance de concepts pour un énoncé donné, et d'autre part, d'un algorithme de gestion du dialogue intégrant des paramètres tels que la tâche en cours, l'historique du dialogue et surtout, le but et le profil de l'utilisateur, de manière à fournir une aide ou une orientation appropriée à la situation.

D'autre part, nous avons observé que certains utilisateurs aimeraient pouvoir interagir avec la machine de manière orale. Actuellement, seul l'ordinateur utilise ce mode de communication, grâce à une synthèse vocale. Nous souhaitons intégrer dans notre système une reconnaissance de la parole, couplée au navigateur de l'utilisateur, et utilisable via le réseau Internet.

La combinaison de ces éléments devrait nous permettre d'affiner notre modélisation et de proposer une véritable application multimodale (clavier/souris/oral) pour une recherche documentaire effective sur le Web.

6 Le système Halpin-Documentaire

"Je crois qu'à la fin de ce siècle, l'usage des mots et l'opinion publique auront tellement évolué, qu'on pourra parler de machines qui pensent sans crainte d'être contredit."

Alan Turing, 1950

Après avoir présenté le système Halpin-Recueil, nous exposons dans cette partie, le système Halpin-Documentaire. Ce dernier a déjà été évoqué dans ce document, mais nous allons ici détailler l'architecture et les composantes orales de dialogues mises en œuvre.

6.1 Présentation du système

Le système Halpin-Documentaire vise à permettre à un utilisateur du World Wide Web de dialoguer avec une machine distante, en langue naturelle écrite ou orale, afin d'effectuer une recherche documentaire. Cela a nécessité la mise en œuvre de composantes de reconnaissance automatique de la parole, de dialogue personne-système, de recherche d'information, et de synthèse vocale.

Les hypothèses de travail que nous posons, pour la conception du système, 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 (aptitudes, handicaps éventuels).

L'utilisateur se trouve, dans Halpin-Documentaire, face à l'écran de l'ordinateur (cf. Figure 25), ce dernier étant 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).



Figure 25 : L'utilisateur dialogue avec le système Halpin-Documentaire.

Nous avons implémenté ce système pour mettre en œuvre une application informatique disponible sur le Web, permettant d'évaluer notre modèle de dialogue Homme-Machine dans des conditions réelles d'utilisation, c'est-à-dire que les usagers ont affaire à un véritable système informatique ayant des capacités dialogiques, et non pas à une expérience du type Magicien d'Oz, dans laquelle un compère se substitue à la machine à l'insu des utilisateurs. Dans Halpin-Documentaire, l'utilisateur peut utiliser deux modalités de recherche d'information : par dialogue écrit ou oral, puis par navigation au sein des hyperliens.

Ce système s'inscrit dans une double perspective d'interaction Homme-Machine : coopérative par le dialogue et réactive en navigation. 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 (sans être guidé). 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 (d'une session à l'autre).

Le système Halpin est essentiellement un *assistant dans la recherche immédiate et dans la collecte*. L'interaction est multimodale, de type exclusive : l'utilisateur tape ou prononce son énoncé, en français, dans une boîte de dialogue. La machine l'analyse et donne une réponse vocale et textuelle. Les stratégies de dialogue s'articulent en fonction des compétences de l'utilisateur constatées (novice ou expert) et de son but (retrouver un document connu ou rechercher un ensemble de documents sur un critère donné).

* * *

Dans le cadre de nos travaux en LN sur le Web, nous faisons l'hypothèse que l'aspect sémantique est une composante forte pour un dialogue coopératif et finalisé, et nous centrons nos efforts sur la reconnaissance de concepts²⁷ au sein du discours. En effet, nous pensons, comme [GERARD et NICOLLE 98] que « *les modèles de dialogue intégrant la manipulation de concepts permettent d'obtenir des DHM réalistes lorsque les tâches à accomplir sont bien identifiées et que le dialogue est coopératif et finalisé* ».

L'analyse effectuée sur chaque énoncé de l'utilisateur sera morphologique, puis sémantique. Quelques modélisations de l'interprétation automatique du langage, notamment à l'oral, émergent depuis quelques années : « *Ainsi, un consensus en faveur de l'approche sémantique semble s'être développé au sein de la communauté de la compréhension de la parole, et cela même parmi ces groupes²⁸ qui, à l'origine, étaient plus favorables à l'approche purement syntaxique. Dans l'approche sémantique globale, la compréhension dépend de la reconnaissance d'îlots de mots au sein de la phrase prononcée.* » [DE MORI 94].

Afin d'expliquer le fonctionnement du modèle de dialogue qu'utilise le système Halpin, nous allons tout d'abord présenter la notion de reconnaissance de concepts dans un énoncé. Ensuite, nous décrirons le fonctionnement global du modèle, puis détaillerons les éléments qui le composent.

²⁷ Un concept est une représentation intellectuelle, à la fois générale et abstraite, d'une catégorie d'objets ou d'idées, conçus par l'esprit ou acquis par lui, et permettant d'organiser les perceptions et les connaissances (ex : le concept du temps).

²⁸ Le groupe du MIT [PEARL 84], et le groupe BBN [PIERACCINI et al. 91].

6.2 Reconnaissance de Concepts

Selon [LE NY 79], la compréhension d'une phrase comporte deux opérations essentielles :

1. la décomposition sémantique des mots lexicaux, qui permet de faire apparaître des prédicats élémentaires ainsi que les arguments qui devront être recherchés dans l'énoncé,
2. le ré-arrangement des arguments autour des prédicats élémentaires pour aboutir au sens global de l'énoncé.

Notre méthode de compréhension d'énoncé s'inspire de cette théorie et se fonde sur une reconnaissance des concepts que nous présentons ci-dessous.

6.2.1 Principe

Après avoir envoyé l'énoncé de l'utilisateur en analyse morphologique, on dispose de la forme canonique de cette phrase. La méthode de reconnaissance de concepts que nous avons mise en œuvre présuppose une intention communicative et guide 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 : *Donner moi* = DEMANDER, *ouvrage* = DOCUMENT, *Baudelaire* = AUTEUR, les autres mots de l'énoncé n'étant pas consommés par l'analyse²⁹. La structure de traits (attributs/valeurs) issue de l'analyse sera donc : UTILISATEUR(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 UTILISATEUR (DEMANDER(DOCUMENT(THEME = Baudelaire))).

Prince rappelle que les mots de connections sont importants en français : « *Contrairement à ce que l'on pourrait penser, l'économie des connecteurs est un mauvais calcul, parce que les connecteurs déterminent correctement les relations existant entre les différents termes employés, alors que l'économie des prédicats peut se faire assez souvent sans dommage.* » [PRINCE 96].

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.

²⁹ Sauf en cas d'ambiguïté, pour essayer d'obtenir une analyse plus fine.

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.

En revanche, 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 de fichiers de concepts, pour construire des hypothèses de structures de traits ;
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*).

Notre base de données de concepts est divisée en différents fichiers, selon le type de concept qu'ils contiennent. En effet, certains concepts sont communs à toutes les situations possibles (acceptation, refus...) et d'autres sont spécifiques à la tâche.

Inspiré des travaux de Brun [BRUN 98], notre module de compréhension d'énoncés utilise un algorithme d'identification de concepts en deux étapes qui tente de comprendre les requêtes des utilisateurs.

Sur la forme canonique de l'énoncé, la première phase de reconnaissance de concepts identifie des mots isolés et les remplace par leur concept (noté xxCONCEPTxx). Par exemple, les mots *livre*, *ouvrage*, *bouquin*, *article*, *acte*, *congrès*, *exemplaire*, *journal*, etc., peuvent être substitués par le concept xxDOCUMENTxx.

La seconde phase consiste à remplacer des groupes de concepts par un concept d'ordre supérieur (noté yyCONCEPTyy). Par exemple, yyUNDOyy remplacera xxANNULERxx xxREQUETExx xxPRECEDENTxx. Que l'utilisateur dise « je veux effacer la requête d'avant » ou « annule la commande précédente », le module de compréhension renverra le concept yyUNDOyy. Ainsi, même si les idées exprimées dans la phrase sont formulées de manière chaotique – et c'est souvent le cas à l'oral – cette méthode est susceptible de faire émerger une idée principale. C'est par exemple le cas avec cette phrase « ouais, c'est bien celle d'avant de requête que je veux supprimer ». Ici encore, et malgré un ordre différent par rapport aux deux exemples précédents, le système trouvera la séquence xxPRECEDENTxx xxREQUETExx xxANNULERxx, et lui substituera le concept supérieur yyUNDOyy.

Cette méthode en deux étapes constitue la base du processus de compréhension que nous détaillons maintenant. La Figure 26 montre les différents concepts structurés sous la forme d'un arbre de concepts.

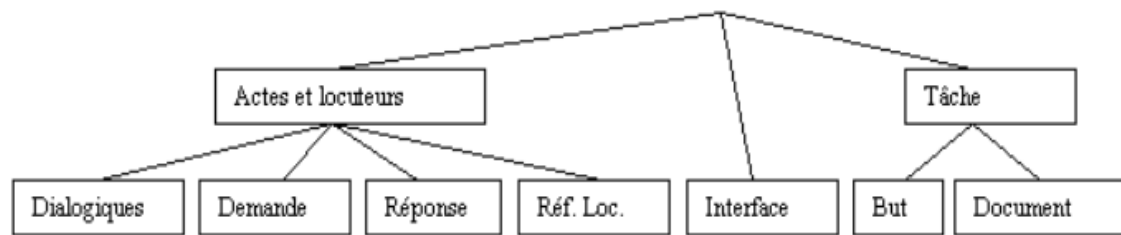


Figure 26 : Arbre de concepts utilisé pour la reconnaissance de concepts dans les énoncés

- Concepts Actes et Locuteurs :

Les concepts Actes et Locuteurs regroupent les actes dialogiques OUVERTURE ; CLOTURE ; POLITESSE ; INSULTE ; AIDE ; les demandes d'action RECHERCHER ; VOIR ; TOUT_VOIR ; AJOUTER ; EXECUTER ; DEMANDER ; ANNULER ; MODIFIER ; les réponses CONFIRMER ; INFIRMER ; les référents aux locuteurs MACHINE ; UTILISATEUR ; 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 ; SYNTHESE ; 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 remonter, à partir d'un énoncé, tous les concepts possibles. On essaye de remplacer un mot ou une séquence de mots par le concept le plus adéquat. Le système travaille d'abord sur des expressions linguistiques, puis sur des mots isolés (Cf. annexe 5 pour voir la trace d'une session Halpin-Documentaire). Par exemple ici pour le concept CONFIRMER, les mots-déclencheurs dans le lexique sont :

en effet ; absolument ; d'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.

Ainsi, la phrase « certes, c'est un rapport original » sera transformée en « xxAFFIRMATIONxx , ce xxETRExx un xxDOCUMENTxx original ». Nous avons choisi de donner priorité aux mots et expressions en relation avec la tâche. Cela n'est pas sans risque et occasionne parfois quelques difficultés. C'est, par exemple, le cas avec la phrase « il n'y a aucun rapport avec ce que j'ai demandé », qui sera transformée en « xxPRONOMxx xxNEGATIONxx y avoir aucun xxDOCUMENTxx avec ce xxQUExx xxUTILISATEURxx avoir xxRECHERCHERxx ».

L'expression « aucun rapport » a été substituée par « aucun document ». Cependant, nous voyons au moins deux interprétations possibles de cet énoncé. La phrase « il n'y a aucun rapport avec ce que j'ai demandé » peut être comprise comme :

- 1) « Existe-t-il des rapports répondant aux critères que j'ai donné précédemment ? »
- 2) « Il n'y a aucune relation entre ce que vous me proposez et ce que j'ai demandé ! »

Dans le cas d'un dialogue oral, une étude prosodique aurait probablement pu nous indiquer si la phrase était interrogative ou non. Dans la négative, c'est la solution « il n'y a aucune relation avec » qui aurait été retenue³⁰.

Cependant, malgré l'absence du point d'interrogation dans l'énoncé original, nous penchons plutôt pour la première interprétation. En effet, au tour de parole suivant, l'usager reformule sa question en demandant « N'as tu aucun rapport intéressant dans ta base de données ? ». Dans cette phrase, il n'y a plus d'ambiguïté : *rapport* est bien employé en référence au « type de document », et non comme synonyme de « relation ».

Cette technique peut paraître assez fruste. En fait sa simplicité est un facteur de robustesse car le module de reconnaissance de la parole engendre beaucoup d'erreurs qui portent surtout sur les mots courts. Seuls les mots assez longs 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.

³⁰ Voici un autre exemple de mauvaise interprétation (cf. le mot « salut ») :

cormoran.lannion.cnet.fr LE Fri Jan 16 11:35:42 MET 1998

11:36:08:Q= est ce que le livre est sur le salut militaire ?

11:36:13:R=Les mots-clés de ce livre sont : Activité de police; Ordre public; Mission policière; Efficacité

11:36:14:R=Ah, enfin un humain qui dit "Bonjour" à une machine.

(...)

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, au regard du corpus ainsi engendré.

6.2.2 Des concepts aux actes

On vient donc d'identifier dans un énoncé quelques concepts possibles. Il s'agit maintenant, pour obtenir du sens, 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 :

UTILISATEUR(CONFIRMER) + UTILISATEUR (REMERCIER) +
UTILISATEUR (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 UTILISATEUR (CONFIRMER) + UTILISATEUR (INFIRMER), mais ne permet pas de traiter les ambiguïtés ou les contradictions comme "oui et non", "oui, mais je ne suis pas d'accord". C'est le gestionnaire de dialogue qui devra gérer ce type de conflit.

6.3 Modèle de dialogue

Nous venons de présenter la notion de reconnaissance de concepts dans un énoncé. Cette notion à elle seule ne suffit pas pour modéliser un DHM coopératif et finalisé.

Le modèle que nous voulons décrire s'inscrit dans une tâche de recherche d'information, au cours de processus interactifs : *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* [BATEMAN et al. 95].

Selon nous, cette situation met en jeu des mécanismes de planification et de gestion des événements qui favorisent une activité de type opportuniste, pour laquelle la machine doit laisser à l'utilisateur un contrôle aussi souple que possible.

La planification intervient dans le sens où pour atteindre ses objectifs, l'utilisateur doit passer par des étapes connues par la machine. Dans notre cas, nous savons qu'il devra fournir des informations, plus ou moins précises, à propos de **l'auteur**, de **l'année** de publication, du **titre**, du **thème**, et du **type** de document recherché. Il n'est pas nécessaire qu'il donne toutes ces informations, mais au moins une d'entre elle devra être fournie pour amorcer la recherche d'information dans la BD.

En ce sens, on peut planifier qu'à un instant T, l'utilisateur utilisera vraisemblablement des mots-clés, des expressions, ou des concepts relatifs aux cinq arguments de recherche cités plus haut.

D'autre part, au cours du dialogue, l'ordinateur devra être en mesure de détecter des intentions de la part de son interlocuteur, compte tenu de la situation et du contexte.

En situation de recherche guidée : la machine se focalise sur un des critères de recherche, et tente de lui donner une valeur en fonction des éléments de réponse produits par l'utilisateur. Par exemple, à la question « Pouvez-vous me donner le nom de l'auteur ? », le système peut s'attendre à une multitude de réponses : oui ; oui, c'est Dupont ; Turing ; non ; pas vraiment ; c'est Boule, mais je ne suis pas certain de l'orthographe ; moi c'est plutôt sur le thème que je voulais travailler ; etc.

En situation de recherche libre : il s'agit ici d'une situation dans laquelle les contraintes sont relâchées. C'est l'utilisateur (expert) qui oriente le dialogue à sa guise, et qui donne des éléments pour remplir ou modifier les *slots* de recherche. Il va par exemple annoncer « Je ne veux que les plus récents ». La machine doit alors détecter la nature de l'information transmise, et déclencher une action de manière à rendre compte de son interprétation de l'énoncé.

Dans un cas, comme dans l'autre, un script sera activé si les conditions nécessaires à son exécution sont remplies. Ce script, sera éventuellement le déclencheur d'un ou de plusieurs autres scripts. Typiquement, lorsque l'utilisateur énonce « cherche avec Durand comme nom », le concept `xxAUTEURxx` émerge, donnant ainsi une précondition au déroulement du script `Recherche_par_Auteur`.

Il faut ensuite détecter si l'un des mots de la phrase est candidat (la majuscule appliquée au nom d'auteur favorise considérablement cette détection). Dans le cas le plus simple, c'est-à-dire s'il n'existe pas de multiples candidatures, l'élément identifié devient un critère de recherche supplémentaire, et le script d'accès à la base de données est déclenché. Dans certaines situations,

des cas particuliers demandent une gestion spéciale : par exemple, la valeur d'un co-auteur ne doit pas venir remplacer la valeur précédente, mais s'y associer.

Plus formellement (cf. annexe 12), il est possible de considérer un dialogue entre A, qui demande de l'information, et B, qui en fournit, de la manière suivante :

Dialogue (A,B) ➔ requête (A,B) + promesse (B, A) + information (B,A) + test-satisfaction (A,B)

ou bien

Dialogue (A,B) ➔ offre (B,A) + acceptation (A, B) + information (B,A) + test-satisfaction (A,B)

Dans le premier cas, le dialogue entre A et B commence par une requête de A à B. Le dialogue se poursuit avec la promesse (de B à A) que l'information va être fournie. Dans un troisième temps, B fournit l'information promise à A. Puis A évalue la pertinence de l'information présentée. Le dialogue peut alors se terminer, ou bien se poursuivre.

Dans le second cas, le dialogue entre A et B commence à l'initiative de B, qui propose d'offrir de l'information à A. Ce dernier peut accepter ou rejeter cette proposition. S'il accepte, B lui fournit cette information, et la suite du dialogue est similaire au premier cas évoqué ci-dessus.

Nous pensons que ce modèle [STEIN et MAIER 95], [STEIN et al. 97], peut être amélioré en intégrant, au sein du processus interactif, des données relatives à l'utilisateur et à ses capacités. Hardie explique que lors de la recherche de documents, les usagers ne cherchent pas tous le même type d'informations : « *some are looking for the ocean and some others for a grain of sand* » [HARDIE 96].

Notre modèle prend donc en compte les **objectifs et le profil de l'utilisateur**, de telle sorte que le système puisse adapter ses réponses selon le profil de l'utilisateur et la tâche en cours. Le modèle que nous proposons permet, pour une tâche définie, finalisée et coopérative de poser la règle suivante :

[Profil] . [But] . [Acte] . [Concepts] . [Tâche] ➔ [Réplique] . [Justification] . [Suggestion]

Pour un profil d'utilisateur donné, un but identifié, un ou des actes de dialogue et des concepts émergents d'un énoncé, et en fonction d'une tâche à réaliser, la machine doit être en mesure de répliquer à cet énoncé, justifier sa réponse si nécessaire, et suggérer des alternatives de manière coopérative pour réorienter le dialogue si nécessaire.

La personnalisation de cette recherche documentaire devrait permettre une plus grande souplesse lors des échanges d'informations personne-système : « *Il s'agit d'acquérir des connaissances sur l'utilisateur de façon directe (en lui posant des questions) ou par inférence (en lui posant des questions indirectes ou par espionnage de son comportement lors d'interactions types) dans le but d'offrir une assistance adaptée au profil de l'utilisateur.* » [SMAÏL 98]. Ainsi, une plus grande pertinence de répliques devrait inciter l'utilisateur à considérer la machine comme un réel partenaire, et à dialoguer avec elle comme avec un humain.

Il en va de même lorsqu'une information est ambiguë ou imprécise, et que l'ordinateur n'arrive pas à extraire de la langue naturelle des éléments pertinents pour lancer une requête appropriée.

En effet, si une phrase est ambiguë, même lorsque le but de l'utilisateur est connu, le système propose des choix. Par exemple, pour les énoncés suivants : « Je cherche un livre d'Alain » et « Donne-moi toutes les thèses d'IHM que tu connais », la machine détecte une ambiguïté liée à la préposition « de », et demande si Alain (respectivement IHM) est bien l'auteur. L'utilisateur confirme ou infirme, et l'accès à la base de données s'effectue selon les bons critères : l'auteur sera Alain dans un cas et le thème sera IHM dans l'autre. Si cet exemple est encore d'un niveau relativement peu ambigu (surtout pour un humain), il n'en va pas de même avec celui-ci : « je souhaite trouver un livre d'algèbre de Boole » pour lequel l'ordinateur devra questionner l'utilisateur pour savoir s'il s'agit d'un document ayant comme thème « algèbre de Boole » (peu importe l'auteur) ou bien un document d'algèbre, écrit par le mathématicien Boole. La première interprétation donne 100 réponses, alors que la seconde en donne 1 seule (sur le serveur de l'INRIA).

Ainsi, nous pensons que, plutôt que d'accéder à la base de données avec une requête incertaine, il vaut mieux résoudre l'ambiguïté en dialoguant avec l'utilisateur.

6.3.1 Gestion du dialogue

Nous avons vu dans les parties précédentes le principe et la mise en œuvre d'une reconnaissance de concepts dans un énoncé en LN. Cette reconnaissance de concepts n'est pas une fin en soi, et s'inscrit dans un processus dynamique plus large qui permet de gérer le DHM de manière relativement efficace.

Le système est chargé de réagir aux différents concepts identifiés, en tenant compte du contexte (ce qu'il est possible de faire à tel moment), du profil de l'utilisateur, de son but déclaré et de l'historique du dialogue.

Le gestionnaire de dialogue (on parle aussi de *dialogueur*) est une partie importante d'un système de DHM. Le dialogueur peut être perçu comme un intermédiaire entre l'utilisateur et l'application [GUYOMARD et al. 84].

On note dans [BILANGE 92] que « *Les fonctionnalités essentielles du module de dialogue sont :*

- *Gérer les tours de parole,*
- *Calculer ce que doit dire le système,*
- *Prédire ce que l'utilisateur est susceptible de dire,*
- *Interpréter les actes de dialogue de l'utilisateur,*
- *Adapter la stratégie de dialogue à la situation.*

Et que les connaissances nécessaires à ce module sont :

- *L'historique du dialogue*
- *Les règles de dialogue,*
- *Les définitions des actes de dialogue. ».*

Nous allons détailler chacun de ces éléments dans le cadre particulier d'un DHM sur le Web, et nous montrerons comment notre gestion du dialogue utilise des notions de planification et de gestions des événements.

6.3.1.1 Gérer les tours de parole

Dans la plupart des dialogues qu'entretiennent des interlocuteurs humains, certaines règles sont respectées. Elles sont souvent sociales ou culturelles, et se résument généralement à la question : qui peut (ou doit) parler ?

Lorsqu'un dialogue s'instaure, c'est souvent l'interlocuteur recherchant de l'information qui engage l'interaction : « Bonjour, avez-vous Hernani de Victor Hugo ? ». Il se peut également que ce soit l'agent détenant de l'information qui déclenche l'échange : « Puis-je vous renseigner ? ».

Généralement, les fonctions sont stables³¹, mais les rôles peuvent être modifiés au cours du dialogue, suivant le contexte et les nouvelles situations qui interviennent.

³¹ Dans un commerce par exemple, un client demeure client tout au long du dialogue, sauf s'il est par ailleurs lui-même commerçant, et qu'il propose, à son tour, de vendre quelque chose à son interlocuteur.

D'autre part, l'éloignement géographique des interlocuteurs accentue souvent le « caractère asynchrone » de la communication, et diminue ainsi la spontanéité des échanges. On peut se couper mutuellement la parole au téléphone, car il est techniquement possible d'écouter et de parler simultanément, alors que c'est impossible sur un forum de discussion d'Internet (chat) ou dans un courrier électronique/postal.

La gestion des tours de parole dans le système Halpin-Documentaire est fondée sur un processus de type talkie-walkie. La machine doit détecter le canal de communication utilisé par l'utilisateur. Il peut s'agir d'une communication écrite, lorsque l'utilisateur se sert du clavier, ou bien orale, lorsqu'il prononce un énoncé dans le microphone.

Une fois le message validé par l'utilisateur, la boîte de dialogue demeure indisponible jusqu'à ce que la machine réplique. Elle est ensuite libérée (et le microphone redevient utilisable) pour l'intervention suivante de l'utilisateur.

Si dans une communication humaine, comme l'a montré Watzlawick, on ne peut pas « ne pas communiquer³² » [WATZLAWICK 80], il n'en va pas de même dans une communication Homme-Machine. En effet, les pannes ou les erreurs entraînent quelquefois des ruptures du DHM voire d'interminables attentes de la part de l'utilisateur, qui se demande si la machine est toujours opérationnelle ou pas. Cela se vérifie davantage sur le réseau Internet. C'est pourquoi nous avons équipé le gestionnaire des tours de parole d'un chronomètre, permettant au système de se rendre compte, après un laps de temps défini, que la liaison sur le Web est perturbée.

La Capture 16 ci-dessous donne un exemple de métadiologue, dans lequel l'ordinateur explique (en M15) pourquoi il ne peut pas traiter correctement ce qui vient d'être dit par l'utilisateur, et lui demande de répéter l'énoncé.

```
194.2.143.1(mailhost.coplanet.fr) SUR Mozilla/4.0 (compatible; MSIE 4.01; Windows 98) LE
8:4:99
(...)
23:28:14:M11=Est-ce que informatique est bien l'auteur du document.
23:28:29:H11=non le thème patate
23:28:47:M12=J'ai trouvé 1659 documents ayant comme thème informatique. Voulez-vous affiner
votre requête ?
```

³² C'est-à-dire que l'on communique toujours (indirectement), même si l'on ne communique pas directement : par exemple, en restant silencieux, je communique quand même mon désir de ne pas parler.

23:29:02:H12=oui
23:29:11:M13=Afin d'affiner, vous devez choisir parmi les critères : Auteur, Titre, Thème, Année, ou Type de document.
23:29:28:H13=comment
23:29:36:M14=Vous ne voulez pas choisir de critère ?
23:29:51:H14=si mais je voudrais un exemple
23:50:00:M15=Il semble y avoir un problème sur le réseau Internet. Je n'ai pas pu récupérer votre énoncé. Répétez-le s'il vous plaît.

Capture 16 : Exemple de métadiologue : l'ordinateur explique pourquoi il interrompt le dialogue relatif à la tâche

De même, lorsque la machine détecte un message vide, elle ne prend pas en compte cette interaction, et informe l'utilisateur avec la phrase « Attention : Vous n'avez rien tapé ». Puis elle redonne la parole à son interlocuteur.

Outre l'historique du dialogue, que l'utilisateur peut consulter à tout moment (avec possibilité de faire des copier/coller vers la boîte de dialogue, pour répéter une phrase par exemple), il est possible avec Halpin de réécouter la dernière intervention de la machine. L'utilisateur utilise cette fonction s'il a été distrait lors de la première énonciation, s'il y a eu des bruits dans la pièce qui ont perturbé son écoute, s'il veut mieux comprendre un message qu'il aurait partiellement assimilé, ou encore lorsqu'il reprend le cours de la conversation, après avoir navigué au sein des hyperliens, et qu'il souhaite se remémorer l'état d'avancement du dialogue.

Enfin, lorsque l'utilisateur valide un énoncé qu'il était en train de taper, la synthèse vocale en cours de prononciation s'interrompt immédiatement. Cet artefact permet à l'utilisateur de couper la parole à la machine (l'expression familière « couper le sifflet » serait d'ailleurs plus appropriée dans ce cas).

6.3.1.2 Calculer les interventions du système

Si la gestion des tours de parole règle la question « quand faut-il intervenir ? », le calcul des interventions du système, quant à lui, est le moyen de traiter la question « que doit dire la machine ? ».

Pour cela, le système dispose de quelques phrases-type qu'il peut utiliser, notamment en début et en fin de dialogue, comme « Dans la page de droite, je vous propose une visite guidée du système Halpin Documentaire. Il suffit de lire, et de cliquer sur les hyperliens. ». Mais, bien entendu, la richesse d'un dialogue émane principalement de la pertinence et de la justesse des informations échangées par les interlocuteurs qui construisent du sens à mesure qu'ils progressent dans leur conversation. La machine, et plus particulièrement le dialogueur, doit être en mesure de calculer des

interventions favorables à la co-construction du sens et appropriées à l'avancement de la tâche en cours.

C'est ainsi que la machine interroge l'utilisateur pour obtenir plus de renseignements, qu'elle lui indique qu'elle n'a pas obtenu ce qu'elle voulait, ou bien encore qu'elle n'a pas compris ce qui a été dit, etc.

De plus, lorsque le système reçoit les résultats des requêtes de la base de données, le dialogueur traite ces informations (conjugaison des verbes, accords en genre et en nombre), de manière à transcrire, en LN, les résultats obtenus.

Dans notre système, la machine ne se contente pas de générer des phrases à trous, du type « j'ai trouvé x documents pour votre requête ». D'une part parce que lorsque x prend pour valeur 0 ou 1, la phrase générée est grammaticalement incorrecte, et d'autre part, parce que les humains accompagnent la plupart du temps un résultat de ce type par un commentaire ou un élément permettant d'enchaîner le tour de dialogue suivant.

Ce peut être une demande de satisfaction (« c'est bien ce que vous cherchiez ? »³³), une proposition coopérative anticipée d'amélioration des résultats (« je vais voir si je ne trouve pas mieux dans des archives plus récentes »), voire une amorce de clôture de dialogue (« vous ne trouverez rien d'autre ici, c'est tout ce que nous avons »). Nous avons incorporé dans Halpin des relances de ce type, qui prennent en compte le but de l'utilisateur, son profil, ainsi que l'état d'avancement du dialogue, pour formuler des répliques originales et adaptées à chaque utilisateur.

M6=Je n'ai pas trouvé de thèse avec comme auteur Miguel Membrado.
M29=J'ai trouvé 13 documents de 1998 ayant comme thème environnement.
M26=Je n'ai pas trouvé de document récent avec comme auteur Ritchie.
M9=J'ai trouvé 14 rapports de 1990 avec comme auteur Pierrel. Voulez-vous affiner votre requête ?
M23=J'ai trouvé 4 documents de 1999 ayant comme thème compression. Consultez-les sur la page de droite. Est-ce que cela vous convient ?
H23=où oui salaire ³⁴ très bien

Tableau 14 : Exemples d'adaptation de la réponse en fonction des critères de recherche

Le Tableau 14 ci-dessus montre que notre système intègre dans ses réponses les connaissances qu'il détient à propos de la tâche en cours. Il précise le type de document s'il le connaît : thèse, rapport, etc. Il utilise une négation pour dire qu'il n'y a pas de réponse à telle requête ; plutôt que de répondre « j'ai trouvé 0 documents ». Le but et le type d'utilisateur sont également pris en compte dans

³³ Extraits du corpus de DHH « Archives départementales de l'Isère ».

les exemples du Tableau 14, mais nous verrons cela plus en détail dans le paragraphe «Adapter la stratégie de dialogue à la situation». L'annexe 9 donne de plus larges extraits de DHM issus du corpus Halpin-Documentaire.

6.3.1.3 Prédire les interventions de l'utilisateur

Étant données la situation courante du dialogue et l'(ou les) intervention(s) précédente(s) de la machine, le dialogueur doit pouvoir prédire quel pourrait être le type d'intervention à venir, de la part de l'utilisateur.

➤ **Planification**

Pour cela, on aborde traditionnellement le problème grâce à une approche par plans : « *L'approche différentielle, ou approche par plans, considère que la langue naturelle est le moyen d'identifier et d'instancier un contexte commun aux deux interlocuteurs, contexte qui conditionne la production des énoncés ultérieurs et qui facilite leur compréhension.* » [GUYOMARD et al. 95].

La planification est donc vue, traditionnellement, comme une activité qui consiste à sélectionner une série d'actions, selon un ordre plus ou moins défini, afin de passer d'un état initial à un état final souhaité.

Dans sa thèse, Lehuen identifie quatre approches distinctes pour une modélisation du DHM : la planification, la structuration, la négociation et l'interaction [LEHUEN 97]. A propos de la planification, il cite un passage de l'article fondateur de James Allen et Raymond Perrault (*Analyzing Intention in Utterances*) : « *Nous partons de l'hypothèse que les agents essayent de reconnaître les plans de leurs interlocuteurs en vue de les utiliser lors de la construction des réponses. [...] Ce modèle est en mesure de rendre compte des réponses qui fournissent davantage d'informations qu'il en est explicitement demandé, des réponses à des fragments de phrases, ainsi qu'à des actes de langage indirects.* » [ALLEN et PERRAULT 80].

Depuis les années 80, plusieurs modélisations, notamment [CARBERRY 90], [ALLEN et LITMAN 87], font référence à la reconnaissance et à la représentation de plans pour la gestion du DHM. Diane Litman, par exemple, fonde son modèle sur trois catégories : plans du domaine, plans du discours et actes de langage [LITMAN 85].

³⁴ Pour l'anecdote, nous avons laissé dans cet exemple la réponse de l'utilisateur : cette intervention était probablement orale et la reconnaissance de la parole a donné « salaire » à la place de « ça à l'air ».

En se basant sur les travaux en compréhension de récits³⁵, la notion de métaplan fut reprise et adaptée à la modélisation du dialogue. On parle de métaplan lorsqu'un plan contient lui-même un ou plusieurs plans comme arguments : « *Un métaplan dont les objectifs seraient d'économiser du temps ou des ressources physiques, pourra, en raisonnant sur ces deux plans, les fusionner pour n'en faire qu'un seul qui ne contiendra qu'une seule instance de ' PAYER' »* [GUYOMARD et al. 95].

Ces modélisations sont principalement basées sur la compréhension des états mentaux des interlocuteurs : ce que l'on *croit* et ce que l'on *sait*, de soi-même et d'autrui. Les croyances et intentions conversationnelles des interlocuteurs ont remarquablement été étudiées par [WATZLAWICK 80], notamment grâce au fameux « dilemme des prisonniers », dans lequel il étudie les comportements sociaux et rationnels des agents (ce que je pense qu'il pense que je pense, etc.).

On comprend bien qu'une machine capable d'appréhender les situations qui se cachent derrière le dialogue aurait une meilleure maîtrise des causes et des effets des propos des interlocuteurs ainsi qu'une meilleure « compréhension » des objectifs qu'ils poursuivent lorsqu'ils parlent. Les humains le font naturellement, comme en témoigne cet extrait de dialogue entendu dans un grand magasin :

Q1 = Vous payez par chèque ou en espèces ?

R1 = (tout en cherchant dans son sac) Je ne sais pas si j'ai pris un stylo ?

Q2 = La machine le remplit toute seule ... vous avez une pièce d'identité ?

R2 = Non mais c'était pour signer ...

En Q2, la caissière reconstruit le plan de son interlocuteur : si elle cherche un stylo, il est donc probable que la cliente souhaite payer par chèque.

En R2, la cliente reconstruit à son tour le plan qu'elle imagine que la caissière s'est construit à son sujet : elle résout la référence relative à l'anaphore et conclut que « le » est employé pour « chèque », puis elle donne implicitement une information supplémentaire : elle sait très bien que la machine remplit le chèque automatiquement. Si l'interprétation humaine arrive facilement à comprendre que le « non » donné en R2 n'est pas la réponse à la question posée en Q2, qu'en sera-t-il pour un ordinateur, s'il n'a pas assez d'éléments pour reconstruire le plan qu'a établi son interlocuteur ?

³⁵ Voir les publications de Wilenski à ce sujet.

➤ **Gestion des événements et des situations.**

A un instant donné du dialogue, le système doit tenter de faire émerger de l'énoncé de l'utilisateur une intention.

Lorsque c'est la machine qui a l'initiative, elle a pour fonction de gérer la situation, et d'amener le dialogue vers des éléments importants pour la tâche choisie. Il s'agit par exemple de déterminer le ou les critères connus par l'utilisateur pour une recherche d'informations.

Étant donnée une situation, l'utilisateur est susceptible de répondre de plusieurs manières différentes. Par exemple, avec la question "Connaissez-vous le nom de l'auteur ?", le système se trouve au niveau du plan "Demander_Auteur" et va réagir en essayant de détecter des événements de manière opportuniste (oui ; non ; c'est Dupont, etc.).

Vilnat, à propos de son système DIABOLO, parle de « *ce que l'utilisateur est susceptible de dire (ce qui implique qu'il n'est pas obligé de le dire), en raison des buts et des plans en cours de développement.* » [VILNAT 97].

Lorsque c'est l'utilisateur qui possède (au moins partiellement) le contrôle du dialogue, il peut plus facilement passer d'une idée à une autre, et cette situation est, selon nous, d'un point de vue informatique, plus difficile à gérer que la précédente. Une intervention du type « il y a beaucoup trop de réponses », devra être interprétée dans un premier temps comme une non satisfaction de l'utilisateur, par rapport aux résultats apportés, et permettra, dans un second temps de supposer quelle pourrait être l'orientation à venir du dialogue. Le nouveau plan est *a priori* connu : l'utilisateur va sans doute vouloir affiner sa requête, mais cette hypothèse n'est pas la seule valide.

Pour l'application Halpin-Documentaire, en se basant sur le corpus Halpin-Recueil et sur des travaux du domaine concernant la même tâche [LEHUEN 97], nous avons recensé un certain nombre d'événements susceptibles d'intervenir au cours du dialogue. La Figure 28 ci-après montre un schéma de la gestion des événements du système Halpin-Documentaire. Ce n'est bien entendu qu'un extrait, et l'ensemble des événements ainsi que des relations entre elles n'y sont pas exhaustivement représentés. Dans certains cas (comme en ouverture de dialogue), les actions possibles sont limitées, tandis que dans d'autres cas (comme en recherche libre, cf. Figure 27), l'utilisateur a beaucoup plus de choix et de possibilités de dialogue avec la machine.

Au début du dialogue, la machine a l'initiative : elle se présente et propose à l'utilisateur de choisir un but parmi trois choix possibles (« Que souhaitez-vous faire ? 1 : Retrouver un document que vous connaissez, 2 : Rechercher un ensemble de documents, ou 3 : Découvrir ce site ? »).

Le système s'attend donc, après une éventuelle salutation, à ce que l'utilisateur réponde à cette sollicitation. Ensuite, la machine cherche à déterminer le profil de l'utilisateur : elle lui propose son aide pour la suite du dialogue. S'il accepte, il est considéré comme un novice, sinon comme un expert. Ce profil d'utilisateur n'est pas fixé une fois pour toutes : il peut évoluer au gré du dialogue, soit par le fait de l'utilisateur lui-même (un expert qui demande de l'aide peut redevenir novice s'il le souhaite), soit du fait de la machine, qui, détectant une série d'échecs dialogiques va proposer son aide à l'utilisateur.

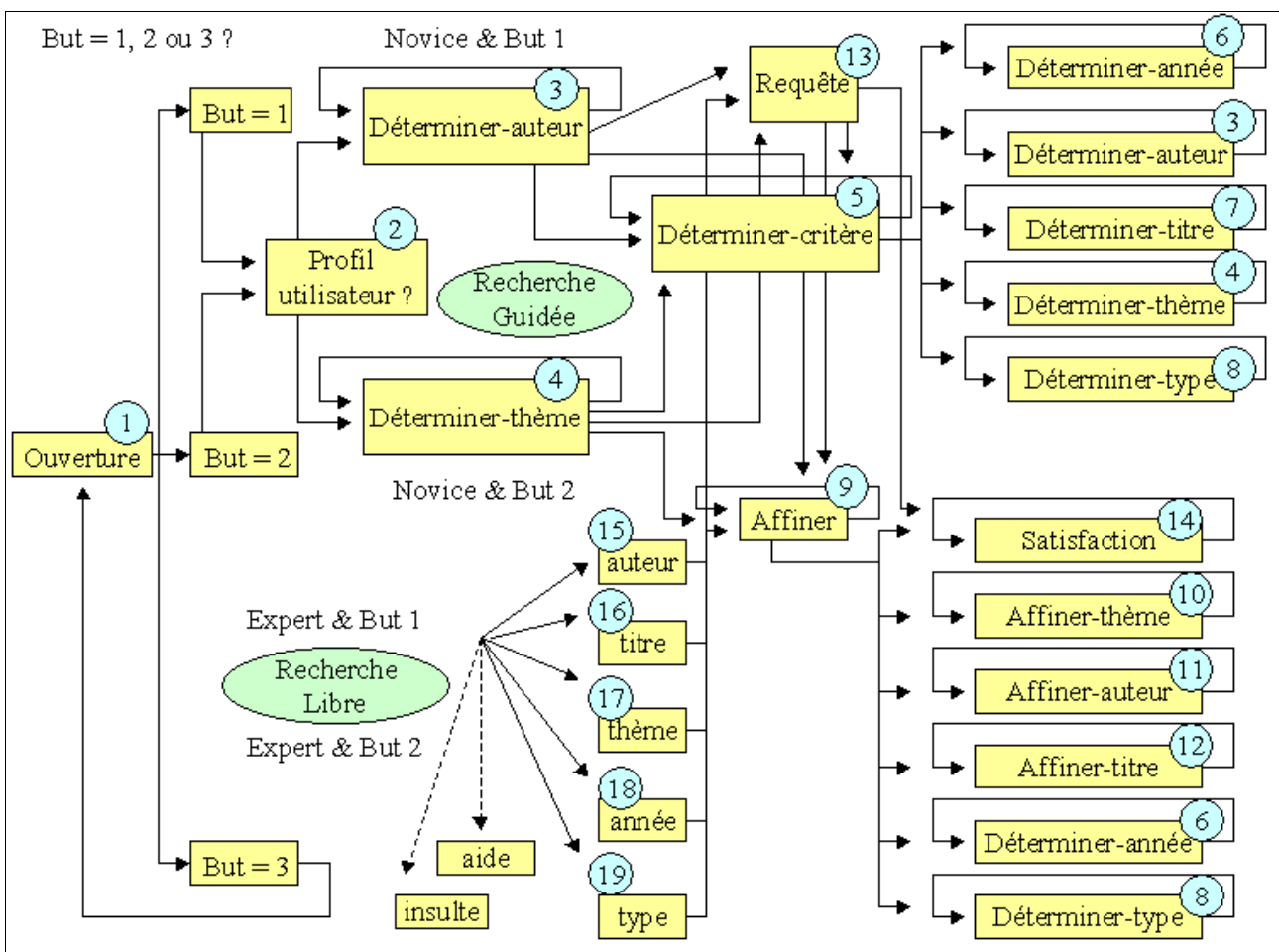


Figure 27 : Gestion des événements du système Halpin-Documentaire

A chaque événement (ils sont numérotés sur la Figure 27), correspond une fiche permettant de réagir à l'énoncé formulé par l'utilisateur.

Prenons l'exemple de l'événement numéro 9 : « Affiner ». Il peut être activé dans plusieurs cas : lorsqu'au cours du dialogue la machine l'aura suggéré à l'utilisateur novice (« J'ai trouvé 1659 documents ayant comme thème informatique. Voulez-vous affiner votre requête ? »), ou encore parce que l'utilisateur expert l'aura demandé (« Peut-on affiner un peu ? »). Dans ce dernier cas, ou

bien lorsque l'utilisateur accepte d'affiner sans donner plus d'informations, la machine rappelle les critères utilisables pour affiner une requête, dans cette base de données (Auteur, Année, Titre, Thème, Type de document).

Le Tableau 15 ci-après explique le traitement de l'événement « Affiner » en fonction de ce qui aura été dit par l'usager, et indique la manière dont doit réagir la machine.

Dans la première colonne, si, à la proposition d'affiner faite par la machine, l'usager donne une réponse négative, le dialogueur doit alors lui demander « Dites-moi si ces résultats vous satisfont », et déclencher l'événement « satisfaction » (n°14 sur la Figure 27) pour analyser le prochain tour de parole.

Réponse Négative	Pas de réponse interprétable	Détection de critères dans la réponse
1. Dire « Dites-moi si ces résultats vous satisfont ». 2. Activer événement « Satisfaction ».	1. Dire « Vous ne voulez pas choisir de critère ? ». 2. Activer événement « Affiner » (bouclage).	A. Si Année : Activer événement « Déterminer-année ». B. Si Type : Activer événement « Déterminer-type ». C. Si Auteur : Activer événement « Affiner-auteur ». D. Si Titre : Activer événement « Affiner-titre ». E. Si Thème : Activer événement « Affiner-thème ».

Tableau 15 : Fiche de traitement de l'événement « Affiner »

Dans la deuxième colonne, si la machine ne détecte pas de réponse à la proposition d'affiner, elle la réitère avec la phrase « Vous ne voulez pas choisir de critère ? », et active l'événement « affiner » (c'est-à-dire que l'on boucle sur cet événement) pour le prochain tour de parole.

Dans la troisième colonne, si le système détecte un concept relatif à l'un des critères de recherche possibles, il active l'événement adéquat. C'est alors ce nouvel événement qui prend le relais (c'est toujours la machine qui a la main). Par exemple, si l'événement n° 9 a activé l'événement n°6 (« Déterminer-année »), cet dernier est chargé de vérifier la validité de l'année donnée par l'usager, de la compléter si nécessaire (94 transformé en 1994, par exemple) ou d'interpréter un équivalent donné en langue naturelle (comme « je ne veux que les documents les plus récents »).

Une réponse est ensuite formulée par la machine. Ce peut être un résultat présentant les réponses données par la base de données, une demande de précision dans le cas où la valeur numérique de l'année n'était pas donnée dans l'énoncé, etc.

On remarque, dans la troisième colonne, que les options A et B renvoient vers des événements de type « Déterminer-valeur », alors que les autres renvoient vers des événements de type « Affiner-valeur ». En effet, on ne peut pas, dans cette base de données, affiner l'année (avec un intervalle de date par exemple) ou le type de document, tandis que l'on peut affiner le nom de l'auteur en ajoutant son prénom, ou en précisant le nom d'un coauteur.

Ce ne sont là que les hypothèses les plus vraisemblables, mais il est bien évident que l'utilisateur pourrait réagir d'une autre manière (demander de l'aide, vouloir grossir la taille de la police de caractère, vociférer une insulte bien sentie, etc.), et le dialogueur doit être en mesure de répondre à ces énoncés.

6.3.1.4 Interpréter les interventions de l'utilisateur

L'interprétation des interventions de l'utilisateur est un point crucial du DHM. Dans notre modélisation, elle consiste à mettre en relation les éléments détectés dans la phase de reconnaissance de concepts, et les hypothèses émises par le dialogueur quant à la suite logique vers laquelle peut se diriger le dialogue.

Si la prédiction de ce qu'est susceptible de dire l'utilisateur concorde avec ce que la reconnaissance de concepts a identifié, l'interprétation de l'intervention ne pose pas de problème particulier. Lorsque la machine demande l'année de publication d'un ouvrage, et qu'elle détecte dans la réponse un élément numérique, il y a de fortes probabilités pour que ce soit effectivement la date attendue. Si de surcroît, un concept `xxDATExx` émerge (ce peut être le mot « date » lui-même, mais également les mots « année », « époque », etc.), cela augmente les chances d'être en présence de l'élément prédit, à moins qu'une ambiguïté ou qu'un élément perturbateur ne soit décelé (comme par exemple la négation dans « je ne veux surtout pas l'année 98 »).

Enfin, lors de la préparation des interventions du système, des règles de dialogue calculent s'il n'y a pas lieu d'interpréter plus efficacement certains énoncés, compte tenu du contexte et de l'historique du dialogue. Pour clarifier ce point, prenons l'exemple d'une règle d'Halpin, qui stipule que si le module de reconnaissance de concepts détecte une « ouverture de dialogue » et que le nombre de tours de parole est supérieur ou égal à trois, cela doit être interprété comme une « clôture de dialogue ». Ainsi, « salut », « tchao », ou « bonsoir » seront assimilés à des formes de clôture de dialogue, si cette règle est vérifiée.

6.3.1.5 Adapter la stratégie de dialogue à la situation

Le système doit être à même d'adapter sa stratégie de dialogue le plus efficacement possible avec les connaissances du monde qu'il possède à un instant donné : cela inclut les informations qu'il reçoit de l'extérieur (d'une base de données par exemple) et celles qu'il détient à propos de l'utilisateur. Selon le contexte, il va adapter ses répliques à la situation, de manière à se rapprocher du but fixé par l'utilisateur.

Le modèle utilise trois stratégies parmi celles décrites dans [CAELEN 92] :

- **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.

Le Tableau 16 ci-après fait apparaître la progression de deux dialogues recueillis avec le système Halpin-Documentaire, le même jour et quasiment à la même heure. Pour le dialogue de gauche, avec un novice, la stratégie de dialogue est coopérative. Pour celui de droite, avec l'expert, elle est essentiellement réactive.

Nous allons analyser comment le dialogue évolue, dans chaque cas, et la manière dont la machine réagit afin d'intégrer dans ses interventions des éléments pertinents, selon le profil et l'objectif de l'utilisateur [ROUILLARD 99].

<i>129.88.41.73 (dracon.imag.fr) SUR Mozilla/4.0 (compatible; MSIE 4.01; Windows NT) LE 14 :8 :99</i>	<i>129.88.41.73 (dracon.imag.fr) SUR Mozilla/4.0 (compatible; MSIE 4.01; Windows NT) LE 14:8:99</i>
18 :43 :34 :M0=Que souhaitez-vous faire ? 1 : Retrouver un document que vous connaissez, 2 : Rechercher un ensemble de documents, ou 3 : Découvrir ce site ?	18:49:19:M0=Que souhaitez-vous faire ? 1 : Retrouver un document que vous connaissez, 2 : Rechercher un ensemble de documents, ou 3 : Découvrir ce site ?
18 :43 :34 :H0=2	18:49:19:H0=1
	18:49:36:M1=Retrouver un document que vous

<p>18 :43 :46 :M1=Rechercher un ensemble de documents. D'accord. Souhaitez-vous que je vous aide à formuler votre requête ? 18 :44 :21 :H1=oui</p>	<p>connaissez. D'accord. Souhaitez-vous que je vous aide à formuler votre requête ? 18:50:17:H1=pas besoin</p>
<p>18 :44 :26 :M2=Pouvez-vous me donner le thème recherché, entre guillemets s'il vous plaît ? 18 :45 :35 :H2=oui, c'est « compression numérique »</p>	<p>18:50:27:M2=Très bien, je vous laisse faire. Appelez à l'aide si vous êtes en difficulté. 18:50:58:H2= je recherche des thèses sur la compression numérique tu vois</p>
<p>18 :45 :48 :M3=J'ai trouvé 23 documents ayant comme thème compression numérique. Voulez-vous affiner votre requête ? 18 :46 :42 :H3=en effet</p>	<p>18:51:08:M3=J'ai trouvé 6 thèses ayant comme thème compression numérique. 18:51:59:H3=quelle est la plus récente ?</p>
<p>18 :46 :55 :M4=Afin d'affiner, vous devez choisir parmi les critères : Auteur, Titre, Thème, Année, ou Type de document. 18 :48 :31 :H4=le type</p>	<p>18:52:10:M4=J'ai trouvé une thèse récente ayant comme thème compression numérique. (...)</p>
<p>18 :48 :43 :M5=Vous pouvez choisir : ouvrages, congrès, thèses, rapports, ou tous les documents. 18 :49 :15 :H5=essayons avec les theses pour voir</p>	<p>18:56:34:H6=je désire affiner le theme à présent</p>
<p>18 :49 :29 :M6=J'ai trouvé 6 thèses ayant comme thème compression numérique. Consultez-les sur la page de droite. Est-ce que cela vous convient ? 18 :50 :01 :H6=non</p>	<p>18:56:59:M7=Le thème actuel est compression numérique. On pourrait y ajouter par exemple : traitement image ; partitionnement ; compression donnée ; image numérique ; codage image ; . Pour ajouter un thème, donnez le entre guillemets, sinon répondez Non pour passer à autre chose. 18:58:54:H7="test en ligne"</p>
<p>18 :50 :07 :M7=Souhaitez-vous plutôt affiner, modifier ou annuler votre requête ? 18 :50 :50 :H7=je voudrais affiner un peu</p>	<p>18:59:06:M8=J'ai trouvé une thèse récente ayant comme thème compression numérique test en ligne. 18:59:51:H8=ok, au revoir</p>
<p>18 :51 :04 :M8=Afin d'affiner, vous devez choisir parmi les critères : Auteur, Titre, Thème, Année, ou Type de document. 18 :51 :46 :H8=avec la date de publication</p>	
<p>18 :51 :55 :M9=Vous pouvez soit me donner une année, soit effacer l'année en cours. 18 :52 :27 :H9=je ne sais pas trop</p>	
<p>18 :52 :40 :M10=D'accord. Souhaitez-vous alors préciser l'auteur, le titre, le thème, l'année ou le type de document ? 18 :54 :32 :H10=l'année</p>	
<p>18 :54 :45 :M11=Vous pouvez soit me donner une année, soit effacer l'année en cours. Ou bien voulez-vous que je recherche les documents les plus récents ? 18 :55 :21 :H11=oui, je veux bien</p>	
<p>18 :55 :51 :M12=J'ai trouvé une thèse récente ayant comme thème compression numérique. Ses mots-clés sont : sûreté fonctionnement ; test en ligne ; tolérance panne ; microprocesseur ; compression donnée. Donnez un thème entre guillemets pour élargir votre recherche ou bien répondez Non pour passer à autre chose. 18 :57 :50 :H12=voyons ce que ca donne avec « test en ligne »</p>	

18 :58 :04 :M13=J'ai trouvé 3 thèses récentes ayant comme thème test en ligne. Consultez-les sur la page de droite. Est-ce que cela vous convient ?	
18 :58 :35 :H13=oui, parfait	
18 :58 :44 :M14=Très bien, je suis heureux d'avoir pu vous aider dans votre recherche.	
19 :00 :11 :H14=allez bye	

Tableau 16 : Adaptation du dialogue selon le contexte et les connaissances de la machine

On observe donc que, pour arriver *grosso modo* au même résultat, le dialogue avec le novice (celui de gauche), est plus long que celui avec l'expert. En est-il pour autant moins efficace ? Nous ne le pensons pas :

- ⇒ En M3, après avoir annoncé le nombre de documents satisfaisant la requête, la machine demande à l'utilisateur s'il souhaite l'affiner. En effet, connaissant son but (cf. H0), et obtenant de la base de données 23 documents, l'ordinateur incite l'utilisateur à donner un autre critère de recherche afin d'obtenir des réponses plus pertinentes.
- ⇒ En M6, la machine se rend compte que le but de l'utilisateur semble être atteint. Elle lui propose de consulter les hyperliens de ces 6 thèses, et cherche à connaître son niveau de satisfaction vis-à-vis des réponses qu'elle vient de fournir.
- ⇒ En M7, comprenant que cela ne convient pas, la machine propose des solutions pour faire avancer le dialogue et pour tenter de répondre au but de l'utilisateur.
- ⇒ En M10, voyant que son interlocuteur « *ne sait pas trop* » quelle date choisir, le système décide de redonner la liste des critères sur lesquels peuvent se faire une recherche.
- ⇒ En M11, cette fois, comme l'utilisateur persiste à vouloir affiner avec la date de publication, et sachant qu'au tour de parole précédent, il n'avait pas donné de valeur pour l'année, l'ordinateur propose une troisième alternative : rechercher les documents les plus récents.
- ⇒ En M12, une seule thèse récente a été retrouvée. Cela n'est plus en adéquation avec le but de l'utilisateur (retrouver un ensemble de documents). C'est pourquoi la machine donne les mots-clés de ce document, pour permettre à l'utilisateur d'élargir sa recherche.

Quant au dialogue avec la personne experte :

- ⇒ En M3, la machine a détecté dans l'énoncé de l'utilisateur le type de documents recherchés. Elle précise donc qu'elle a trouvé 6 thèses en réponse à cette requête. Notons au passage qu'elle n'a pas effectué une recherche sur « *la compression numérique tu vois* », mais qu'elle a su, grâce à la reconnaissance de concepts, distinguer les mots qui allaient servir à formuler la requête.

- ⇒ En M3 toujours, même si le but de l'utilisateur n'est pas atteint, la machine n'alourdit pas le dialogue (avec des propositions pour affiner la requête), et laisse l'utilisateur libre d'orienter la suite des échanges.
- ⇒ De même, en M8, l'objectif est atteint, mais la machine ne pose pas de question pour savoir si l'utilisateur est satisfait (car c'est un expert).

Nous donnons ci-dessous un extrait de l'algorithme de gestion du dialogue dans Halpin-Documentaire, qui montre comment la machine gère le dialogue en fonction des concepts détectés, des buts ainsi que des actes de dialogue (HOMME est abrégé par H et MACHINE par M) :

Début-Dialogue

Début Ouverture

Affichage de la page d'accueil Web

Stratégie = Directive

Sous-dialogue But (G)

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(Infirmier) ; 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(Infirmier) ; 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).

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

* * *

Après avoir décrit notre modèle de dialogue, nous allons maintenant présenter l'architecture du système développée sur la base de ce modèle.

6.4 Architecture du système

Le système réalisé s'appuie sur un ensemble de modules disposés sur différents serveurs distants communiquant sous le protocole IP (Internet Protocol). On utilise des logiciels du commerce, des logiciels libres (freeware) et des logiciels spécialement développés pour Halpin. L'application finale fonctionne avec un navigateur sous environnement Unix, PC, Mac, ou tout autre système gérant les applets Java. Le dispositif audiophonique (microphone et enceintes) est recommandé, mais n'est pas indispensable au bon déroulement du dialogue entre l'homme et la machine.

Les langages informatiques que nous avons utilisés sont : Java, Perl et HTML pour le développement des parties clientes et serveur, et le langage C pour la réalisation du module d'interface entre le navigateur et le logiciel de reconnaissance de la parole (ViaVoice). Le noyau du serveur représente, à lui seul, plus de 6000 lignes de code Perl. Il gère tous les appels sur le réseau Internet et permet à plusieurs utilisateurs de se connecter simultanément au système. La Figure 28 ci-après présente l'architecture générale du système Halpin-Documentaire.

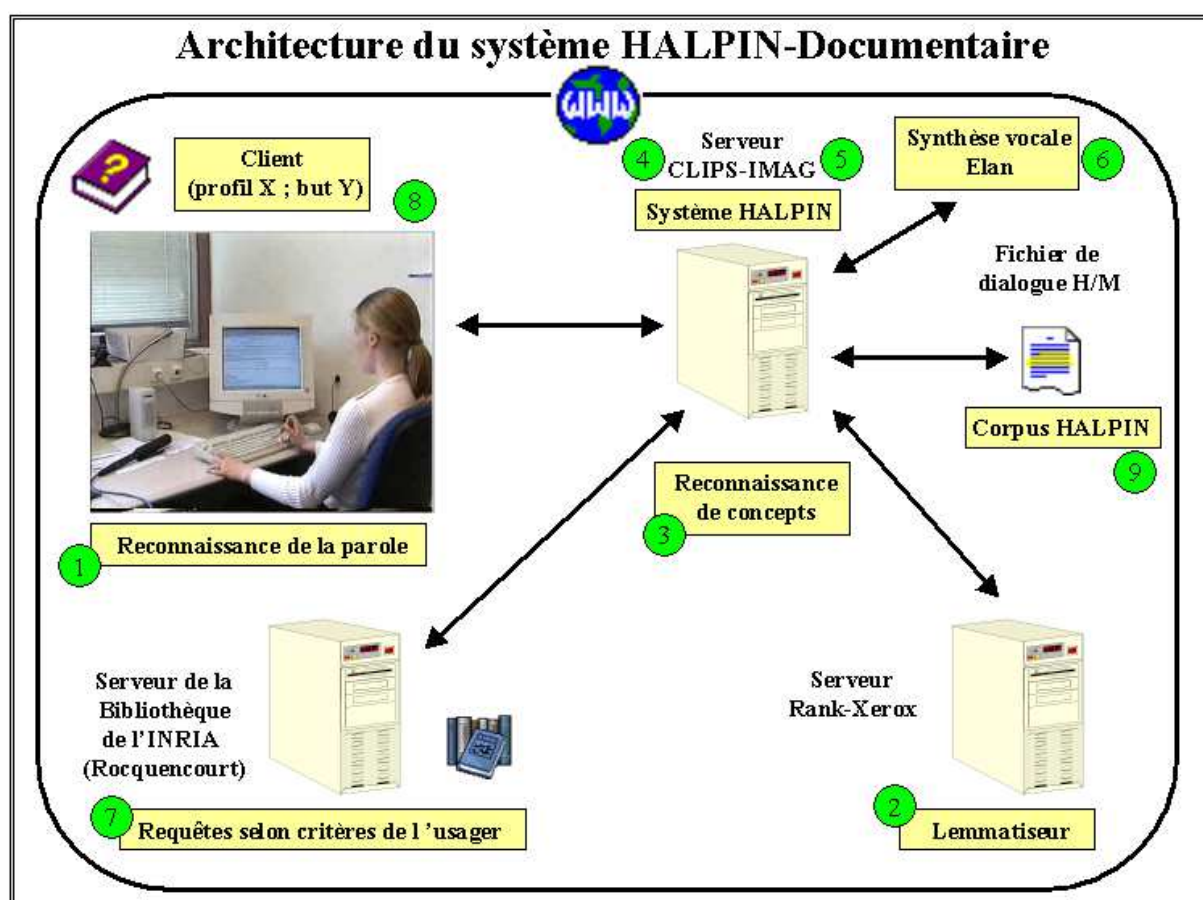


Figure 28 : Architecture du système Halpin-Documentaire

Description des modules :

1. *La reconnaissance de la parole* : c'est le système ViaVoice d'IBM. Un module couplant le logiciel de reconnaissance de la parole au navigateur 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.
2. *L'analyseur morphologique* : c'est celui du Centre de Recherches de Xérox [GAUSSIER et al. 97] auquel on accède de manière distante. Ce module délivre la forme lemmatisée de chacun des mots de la source d'entrée.
3. *La compréhension par concepts* est un module développé dans le cadre d'Halpin [ROUILLARD 98], [ROUILLARD et CAELEN 98]. Il fournit une structure de traits à partir d'une analyse conceptuelle de la source lemmatisée. Ce module est exécuté sur un serveur de notre équipe (au laboratoire Clips de Grenoble).
4. *Le contrôleur de dialogue* a été également développé dans le cadre d'Halpin [ROUILLARD 99]. Il permet de gérer les tours de parole et l'avancée dans la tâche.
5. *La génération des énoncés* de réponse est un module qui s'appuie sur un modèle de génération à "trous", mais augmenté par la prise en compte de l'historique des échanges, de l'objectif et des capacités de l'utilisateur.
6. *La synthèse vocale* est le système de la société française Elan informatique. Ce module fonctionne en accès distant, et demeure sur notre serveur. Contrairement à la reconnaissance de la parole, l'utilisateur n'a pas besoin de l'installer sur sa propre machine.
7. *L'accès à la base de données* documentaire de l'INRIA se fait de manière distante sur le serveur dédié de l'INRIA. Les pages réponses récupérées au format HTML sont analysées et traitées, pour être envoyées au dialogueur du système Halpin.
8. *L'interface graphique* et la gestion des dispositifs s'exécutent sur la machine cliente. Ce module tourne dans une applet Java sous un browser comme Internet Explorer ou Netscape.
9. *Le recueil du corpus HALPIN* s'effectue au fur et à mesure des connexions, de manière anonyme. L'utilisateur peut également donner son avis sur le système, en remplissant un formulaire de satisfaction, en ligne (cf. partie 6.6.1).

On constate donc que le système Halpin-Documentaire fonctionne avec la même architecture que le système Halpin-Recueil, présentée au chapitre 5. Les différences notables sont :

1. l'accès à une base de données spécialisée dans le domaine de l'informatique et de l'automatique (INRIA). On n'utilise donc plus le serveur (géographiquement proche de notre laboratoire) de la bibliothèque de l'université Pierre Mendès France, à Grenoble, mais celui de l'INRIA, situé à Rocquencourt, près de Paris.
2. une synthèse vocale de meilleure qualité que l'on a intégrée dans une applet Java, et que l'utilisateur peut arrêter ou réécouter à volonté.
3. une reconnaissance de la parole, permettant une interaction orale avec le système Halpin-Documentaire, de manière directe ou pas. Dans ce second cas, l'utilisateur peut corriger un énoncé mal reconnu, ajouter des noms propres qui ne sont pas dans le vocabulaire, etc.

6.5 Composantes orales de DHM

Notre objectif est de permettre à un usager du World Wide Web de dialoguer de manière orale, en langue naturelle, avec un ordinateur distant (notre serveur), afin d'effectuer une recherche d'informations documentaires. Nous avons choisi ce domaine parce qu'il nous paraît utile de proposer une approche dans laquelle l'utilisateur trouve réellement une motivation à employer le système Halpin-Documentaire que nous lui soumettons. Les expériences en laboratoire sont parfois biaisées par le fait que les pseudo-utilisateurs (souvent des collègues de travail) n'utilisent pas les logiciels qu'ils testent comme ils le feraient en situation concrète et réelle. Par ailleurs, les résultats obtenus sont souvent faussés à cause du manque de réalisme de la tâche à réaliser. Dans notre cas, la tâche à accomplir est finalisée, réaliste et motivante. En contrepartie, les utilisateurs seront d'autant plus critiques vis-à-vis du système proposé qu'ils savent relativement bien obtenir le même genre d'informations (des références documentaires) par des moyens traditionnels (en interrogeant un(e) bibliothécaire notamment). Malgré les problèmes techniques du réseau Internet, comme sa lenteur³⁶ et son encombrement, qui sont de sérieux handicaps de départ pour notre projet, le Web, outil interactif et multimédia s'il en est, nous paraît être un terrain propice et très intéressant pour nos expériences de DHM.

³⁶ Certains utilisateurs donnent une autre signification à WWW : Wait Wait Wait

Cependant, notre dessein est confronté à deux principaux écueils :

(1) **L'aspect technique** : comment donner au système des facultés de dialogue via le réseau Internet. Cela se traduit par :

- Comment communiquer oralement avec la machine : quels outils de reconnaissance de la parole, dans quelles conditions (hésitations, reprises, etc.), et avec quelle architecture (reconnaissance locale ou distante) ?

- Comment donner une voix à la machine : quels outils de synthèse vocale : comment les intégrer dans les navigateurs standards actuels pour une couverture logicielle maximum (c'est-à-dire multiplateforme) ?

(2) **L'aspect dialogique** : comprendre ce que dit l'utilisateur. Cela revient, après la phase acoustico-phonétique précédente, à comprendre le sens des propos de l'utilisateur, ses buts, sa stratégie de recherche, ainsi que les actions qu'il cherche à faire faire à la machine pour arriver à ses fins.

Nous avons discuté l'aspect dialogique dans la partie 6.3 de ce document, et nous nous intéressons à présent, à l'aspect technique de la mise en œuvre de composantes orales de DHM, sur le réseau Internet.

6.5.1 La synthèse vocale

"Ce serait certes une invention considérable, que celle d'une machine capable de reproduire nos paroles, avec leurs sons et leurs articulations.

Je crois que la chose n'est pas impossible. "

Leonhard Euler, 1761.

La synthèse de la parole, ou synthèse vocale, est un ensemble de techniques informatiques et linguistiques permettant de faire prononcer un texte par une machine. Le premier système électrique capable de reproduire la parole humaine fut mis au point au laboratoire Bell, en 1939, par Homer Dudley. Cette technologie ne doit pas être confondue avec des systèmes comme l'horloge parlante, qui assemble des mots ou des phrases préenregistrés par un locuteur humain. Dès les années 70, les handicapés moteurs et sensoriels ont été des utilisateurs intensifs des technologies de synthèse vocale, malgré les imperfections d'alors. Aujourd'hui, avec l'avènement des outils multimédia, ces technologies commencent à être employées dans quelques domaines (aide à la navigation dans les véhicules haut de gamme, service bancaire téléphonique par exemple).

La synthèse vocale d'un texte s'effectue en plusieurs phases : l'analyse syntaxique du texte, la transcription des mots en phonèmes, puis la génération des sons et des mots correspondants, par un synthétiseur, avec les intonations appropriées. Les phonèmes sont les plus petits éléments constitutifs d'un langage. Ils n'ont pas de sens particulier mais possèdent des traits caractéristiques qui permettent de distinguer les mots les uns des autres. Leur assemblage forme des mots. Par exemple, dans le mot « brin », il y a trois phonèmes notés [b] [ʀ] [ɛ̃].

Le nombre de phonèmes est variable d'une langue à l'autre. Il en existe 37 pour la langue française. Ils peuvent être différents selon les langues. Le français ne connaît pas le « th » anglais ni le « j » espagnol, par exemple.

Dans notre projet, nous avons utilisé deux logiciels de synthèse vocale. Celui de la faculté polytechnique de Mons, en Belgique (Mbrola) [DUTOIT et al. 96], [DUTOIT et PAGEL 96], puis celui de la société Élan Informatique [ELAN], fondé sur une technologie mise au point au Centre National d'Études des Télécommunications (C.N.E.T.).

Dans les lignes qui suivent, nous présentons la synthèse vocale comme modalité d'interaction avec un utilisateur du Web. Dans cet échange entre l'ordinateur et la personne connectée, il est indispensable, pour obtenir un dialogue crédible et pertinent, que les répliques orales de la machine s'inscrivent dans un processus dynamique et qu'elles contiennent des éléments calculés à la volée. En effet, le but n'est pas ici d'envoyer via le Web des phrases typiques pré-enregistrées, mais bien de réagir à ce que dit l'utilisateur ; d'intégrer dans la réponse des éléments comme des noms d'auteurs, des dates, des titres, ou le nombre de réponses fournies par la base de données.

Le problème principal dans cette démarche réside dans le fait qu'à la base, le langage HTML utilisé pour le Web est statique, et non pas dynamique. C'est-à-dire qu'en règle générale, une page Web est créée une fois pour toute par son concepteur. Chaque visiteur obtient la même page, contenant les mêmes images, les mêmes ancres (liens hypertextuels) et les mêmes données à chaque connexion. Cette approche traditionnelle n'est pas envisageable dans le cadre du DHM écrit ou oral sur Internet, puisque le dialogue est un processus dynamique. Nous avons donc utilisé le langage Perl afin de programmer un code générant à son tour du code HTML. Il est ainsi possible d'insérer des informations dynamiques dans une page Web, et de faire en sorte que chaque utilisateur connecté entretienne un dialogue personnalisé avec l'ordinateur.

D'autres problèmes techniques que nous avons dû résoudre, comme la gestion du *cache*³⁷ et des *proxys*, ou le manque de compatibilité de certaines machines virtuelles Java ne seront pas évoqués dans ce document.

Pour passer d'un texte écrit à son équivalent sonore, quatre étapes étaient nécessaires dans le système Halpin-Recueil : la génération des phonèmes à partir du texte écrit, la génération de la synthèse vocale avec Mbrola à partir de ces phonèmes, puis la conversion du fichier (format « wav » de Microsoft au format « au » de Sun) et enfin la compression du fichier sonore (cf. annexe 7). On réduit ainsi le poids du fichier final de moitié, pour en faciliter le débit sur Internet, sans trop altérer la qualité du signal que va entendre l'utilisateur. C'est l'équivalent de la parole transmise par le téléphone. Le fichier sonore créé peut alors être encapsulé dans la page Web que va lire le navigateur de l'utilisateur. On obtient un élément sonore qui se lance automatiquement lorsque la page Web est complètement chargée par le navigateur.

Cette première solution fonctionnait bien, mais ne permettait pas de réécouter le son ainsi diffusé par le navigateur. En effet, contrairement à du texte ou une image qui, une fois présentés à l'utilisateur restent disponibles sur la page Web chargée, le son, éphémère, n'était (ici) joué qu'une seule fois. L'efficacité d'une telle interaction avec l'utilisateur peut alors être largement diminuée si les conditions d'écoute ne sont pas optimales : niveau sonore trop bas, bruit dans la pièce, utilisateur distrait, inattentif, surpris, sont autant de risques de voir échouer le dispositif.

Bien que notre système multimodal fournisse, en même temps que la synthèse vocale, le texte correspondant à l'énoncé, ainsi que d'éventuels hyperliens, ce problème ergonomique de manque de contrôle sur le son a été dénoncé par des usagers du Web, lors de l'évaluation, en condition réelle d'utilisation. C'est pourquoi, au cours de notre cycle de développement incrémental, nous avons remplacé la solution présentée ci-dessus par un fonctionnement avec une applet Java, équipée de boutons permettant d'arrêter ou de réécouter les fichiers son, joués par le navigateur.

Les utilisateurs ont également critiqué la qualité de la synthèse vocale (Mbrola) que nous utilisons. Ils dénonçaient la prosodie de cette voix synthétique, monocorde et trop « mécanique » à leur goût. Pourtant, selon nous, le fait que la machine parle avec une voix mécanique n'est pas nécessairement

³⁷ Le but d'un cache ou d'un proxy est d'éviter de rapatrier une page Web si celle-ci a déjà été visitée. Cela pose problème lorsque l'on souhaite mettre en place un processus dynamique, et que le système estime qu'il

un défaut du système. En effet, si l'on observe le DHM d'un point de vue cognitif, cette voix de robot rappelle continuellement à l'utilisateur que c'est à une machine qu'il a affaire. Le fait que l'ordinateur parle de cette manière permet à l'utilisateur de bien prendre conscience du statut et du rôle de son interlocuteur.

En revanche, une voix synthétique qui se rapproche de la voix humaine peut inciter l'utilisateur à attribuer à la machine un niveau de compréhension dialogique en rapport avec son élocution. La personne qui entend une machine parler presque comme un humain, risque d'attendre d'elle des capacités de compréhension proches de celle d'un humain. On tolérera donc moins d'erreurs de la part d'une machine qui parle convenablement, puisque l'on s'imagine que sa façon de s'exprimer est étroitement liée à ses (hautes) capacités calculatoires.

Il faut donc noter que cette évolution vers une meilleure synthèse vocale n'est pas forcément un avantage pour le DHM dans son ensemble. Cependant, nous avons voulu remplacer la synthèse vocale du système Halpin-Documentaire par une autre, de meilleure qualité, plus audible, plus rapidement calculée et plus robuste aux multiples appels en parallèle sur le Web.

Pour cela, nous avons utilisé le logiciel Text To Speech (TTS) Engine, de la société Élan Informatique. C'est un moteur logiciel de synthèse de parole multilingue³⁸ générée à partir du texte. Un jeu de fonctions permet de contrôler totalement la voix, le timbre, le volume, la vitesse d'élocution et les pauses. Lors du chargement du programme en mémoire, il est possible de choisir une voix masculine ou féminine. L'annexe 8 donne les détails techniques du codage (client et serveur) de la synthèse vocale d'Élan dans le système Halpin-Documentaire.

6.5.2 La reconnaissance de la parole

A l'inverse de la synthèse vocale, la reconnaissance de la parole consiste à identifier dans un énoncé prononcé par un locuteur humain, les différents phonèmes qui le composent, puis à retranscrire en mots et phrases écrites ce qui a été dit.

Nous pensons que pour la reconnaissance de la parole sur le World Wide Web, deux solutions différentes sont envisageables : une locale et une distante.

n'a pas besoin de donner une nouvelle version de cette page puisqu'il croit détenir la dernière version, qui en réalité ne l'est pas.

³⁸ Français, anglais, américain, allemand, espagnol ...

Nous présentons ci-dessous les deux méthodes et détaillons le module programmé pour notre projet, reliant le logiciel de reconnaissance IBM ViaVoice à notre système Halpin-Documentaire.

6.5.2.1 Reconnaissance locale

La première solution nécessite que chaque utilisateur installe un logiciel spécifique de reconnaissance de la parole sur son propre ordinateur. Il faut alors utiliser un programme intermédiaire, dédié à jouer le rôle d'interface entre le logiciel de reconnaissance vocale et celui de l'application cible (Halpin dans notre cas).

- Les *avantages* de cette méthode sont : la rapidité de la reconnaissance vocale et le fait que la personne puisse entraîner le système de reconnaissance avec sa propre voix, pour obtenir de meilleures performances.
- Les *inconvenients* sont : la nécessité d'installer un logiciel sur chaque poste client (environ 1000 francs la licence), et la limite du vocabulaire compris par le logiciel standard. En outre, la machine utilisée doit être cadencée à 166 Mhz ou plus, posséder 105 Mo d'espace libre sur son disque dur, un lecteur de CD-ROM pour l'installation du logiciel, 32 Mo de mémoire vive pour un Windows 95 et 48 Mo pour un Windows NT.

Dans notre application, l'interaction est multimodale et l'utilisateur a la possibilité d'entrer son énoncé grâce à l'interface traditionnelle dans la boîte de dialogue (clavier/souris) ou bien il peut le faire de manière vocale, s'il est équipé du matériel audio nécessaire. Cette multimodalité est de type exclusive [NIGAY 94], c'est-à-dire qu'il n'y a pas de fusion entre les informations provenant des différents canaux de communication et que l'usage des modalités s'effectue en séquence. En effet, le DHM n'est guère propice (en entrée) à des actions multimodales synergiques de type « mets ça là » [BOLT 80], qui sont plus utilisées dans des systèmes multimodaux de manipulations graphiques, dans lesquels on peut nommer l'objet vocalement, en même temps que l'on désigne un emplacement avec la souris, par exemple.

Pour notre projet, nous avons opté pour une solution dite « push-to-talk ». On clique sur un bouton, et on le maintient appuyé tout en parlant. On lâche ensuite le bouton pour arrêter. Il est techniquement possible de programmer un système de reconnaissance de la parole avec dispositif de prise de son dit « main libre » basé sur les seuils d'énergie de la voix (cf. Figure 29 – empruntée à

D. Vaufreydaz) : l'enregistrement de la voix démarre au moment où l'on parle³⁹ et se termine après un temps de pause déterminé.

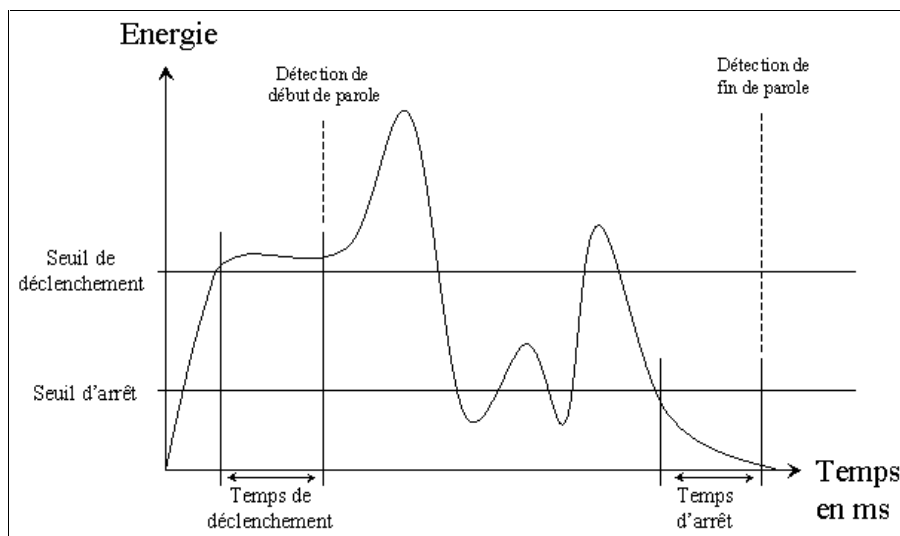


Figure 29 : Détection de parole en dispositif « main libre »

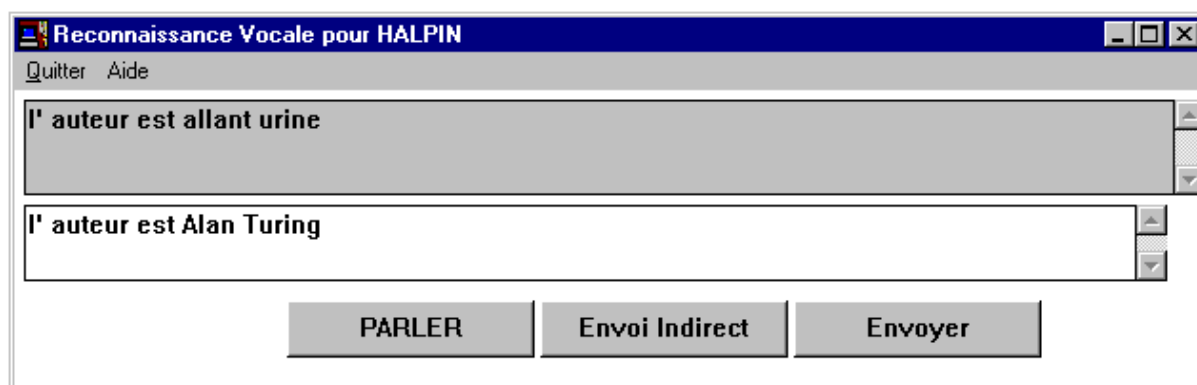
Le module que nous⁴⁰ avons écrit en langage C, est un fichier de 96 Ko (Reco_Halpin.exe) qui nécessite, pour fonctionner, une DLL⁴¹ de 148 Ko (smapi.dll). Ce module est couplé à ViaVoice, et fonctionne avec un ordinateur PC équipé d'un microphone.

La Capture 17, ci-après, est une vue du module de reconnaissance pour Halpin-Documentaire. Dans cet exemple, l'utilisateur a prononcé la phrase « L'auteur est Alan Turing ». La fenêtre grise est celle où s'affiche le texte obtenu grâce à la reconnaissance de la parole. La blanche est une zone dans laquelle on peut éditer, corriger et modifier le texte, si besoin. En mode d'envoi indirect, on voit que l'utilisateur a pu modifier la phrase (mal) reconnue par le logiciel de reconnaissance, car ce dernier ne possédait pas dans son vocabulaire le nom de l'auteur prononcé. Le bouton « envoyer » permet ensuite d'expédier la phrase corrigée (ou pas) vers le serveur qui héberge le système Halpin.

³⁹ En réalité, grâce à une technique de *buffer* (une mémoire tampon), le système écoute et enregistre en permanence le signal émis. Ainsi, il est possible de récupérer le signal, non pas à partir du moment précis où le locuteur commence à s'exprimer, mais un peu en amont, pour ne pas risquer de manquer le début de la phrase.

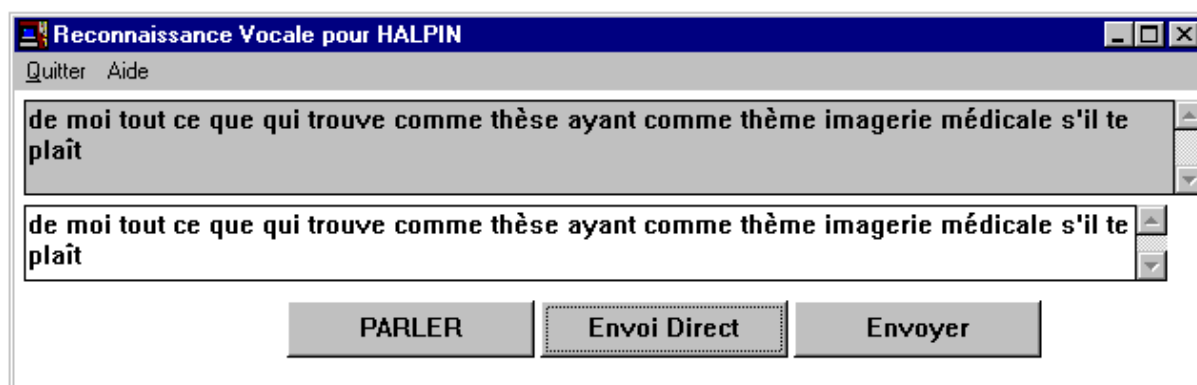
⁴⁰ Mohammad Akbar, Harouna Kabré et moi-même.

⁴¹ Dynamic Link Library.



Capture 17 : Reconnaissance de la parole avec envoi indirect

Sur la Capture 18, ci-après, l'utilisateur a choisi de travailler en mode d'envoi direct. C'est-à-dire qu'il ne souhaite pas modifier ce que le système a reconnu : la phrase, qu'elle comporte des erreurs ou non, est envoyée telle quelle au système Halpin-Documentaire. C'est plus rapide et cela demande moins de manipulations, mais en contrepartie, le risque d'envoi de messages erronés est plus grand. Dans notre exemple, la phrase prononcée était « Donne-moi tout ce que tu trouves comme thèses ayant comme thème l'imagerie médicale s'il te plaît ». Nous voyons qu'ici, les quelques erreurs de reconnaissance n'ont que peu d'influence sur le sens de la phrase, qui demeure tout de même compréhensible.



Capture 18 : Reconnaissance de la parole avec envoi direct

Techniquement, deux problèmes épineux ont dû être élucidés pour permettre à un usager du Web d'utiliser Halpin avec une entrée vocale utilisant ViaVoice.

L'architecture clients-serveur du Web implique qu'un serveur doit être capable de gérer simultanément une multitude de liaisons avec des clients différents. Cela ne pose aucun problème en utilisation traditionnelle lorsque le serveur répond aux demandes d'un navigateur identifié. Mais dans notre cas, le module spécialement développé pour le système d'entrée vocale est indépendant

de tout navigateur. Une fois reconnue, la phrase de l'utilisateur est disponible sous la forme d'une chaîne de caractères, mais demeure dans la machine cliente.

Dès lors, le premier problème consiste donc à trouver un moyen pour transmettre au serveur le texte résultant de la reconnaissance de la parole, et de prendre en compte cette interaction verbale comme si elle avait été tapée dans une boîte de dialogue classique.

Des problèmes techniques d'incompatibilité de versions de langages de programmation (C et C++) ne nous ont pas permis de mettre au point une solution de communication par sockets⁴² directes (c'est-à-dire sans création de fichiers de communication). A défaut, nous avons donc dû élaborer une solution par transmission des données dans un fichier.

Le deuxième problème est alors celui de l'identification des fichiers envoyés, afin que le serveur sache à quel client chaque intervention orale « appartient » ; sans cela, la phrase prononcée par une personne est prise en compte par tous les navigateurs des clients connectés au même moment.

Nous avons choisi d'utiliser, comme nom de fichier, l'adresse IP⁴³ de la machine cliente. Un fichier du type « 129.88.41.73.8.mes » est envoyé au serveur et contient un message (la phrase reconnue). Le serveur est donc en pseudo attente active⁴⁴ d'un message qu'un client déposerait dans un répertoire spécifique. Pour éviter que le serveur n'obtienne qu'une partie du message (dans l'hypothèse où le logiciel client n'a pas encore fini d'écrire le message lorsque le serveur commence à le lire), la lecture du fichier message n'est autorisée qu'une fois l'existence d'un fichier de contrôle⁴⁵ vérifiée (« 129.88.41.73.8.ok » dans notre exemple).

6.5.2.2 Reconnaissance distante

La deuxième possibilité consiste à utiliser un serveur de reconnaissance de la parole unique, auquel les différents clients envoient leur signal acoustique pour traitement à distance. La principale

⁴² Mécanisme de communication utilisé en programmation d'applications réseaux. C'est un port logique établi à chaque connexion via Internet vers un serveur. La "socket" fait apparaître le réseau comme un fichier que l'on peut écrire ou lire.

⁴³ Adresse codée sur 32 bits selon le protocole Internet et affectée à un ordinateur figurant dans un réseau. Une portion de l'adresse IP désigne le réseau et l'autre désigne un ordinateur dans ce réseau.

⁴⁴ En fait, nous avons programmé des temps de pause (sleep) pour que le serveur ne scrute pas en permanence le répertoire convenu. Il le fait régulièrement, mais pas en continu.

⁴⁵ Les anglo-saxons parlent de « flag » (drapeau) donnant un signal d'autorisation.

différence avec une reconnaissance locale réside donc dans le fait que ce n'est plus un fichier texte qui transite sur le réseau Internet, mais un fichier de signal sonore.

Un logiciel doit être installé sur la machine cliente. Il permet de capturer les phrases prononcées par l'utilisateur et de les envoyer au serveur via le réseau. Là encore, un dispositif « main libre » est tout à fait envisageable.

- Les *avantages* de cette méthode sont : d'une part le fait que l'utilisateur n'a pas besoin de posséder un logiciel de reconnaissance de la parole, et d'autre part, le fait que le serveur de reconnaissance soit multilocuteurs, disposant d'un large vocabulaire. Enfin, le fait que la mise à jour du système avec une nouvelle version soit transparente pour l'usager, est un atout supplémentaire. Les coûts de fonctionnement sont moindres : on ne remplace qu'une seule version du logiciel (celle du serveur, pendant la nuit par exemple), et l'on n'est pas obligé de distribuer autant de logiciels qu'il existe de clients.
- Les *inconvénients* sont : de possibles variations selon l'hétérogénéité des microphones utilisés par les clients, et la manière dont ils s'en servent (proximité du micro par rapport à la bouche, utilisation de pauses ou pas, niveau d'articulation, etc.). Un besoin plus important du débit du réseau, puisque l'on envoie un fichier sonore, plus volumineux que son équivalent au format texte. Il faudra certainement utiliser des logiciels de compression pour optimiser ce type de transmission.

6.6 Interface multimodale du système

La Figure 30 ci-après présente l'interface du système Halpin-Documentaire. Elle 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 de navigateur (*browser*) classique, comme Internet Explorer de Microsoft.

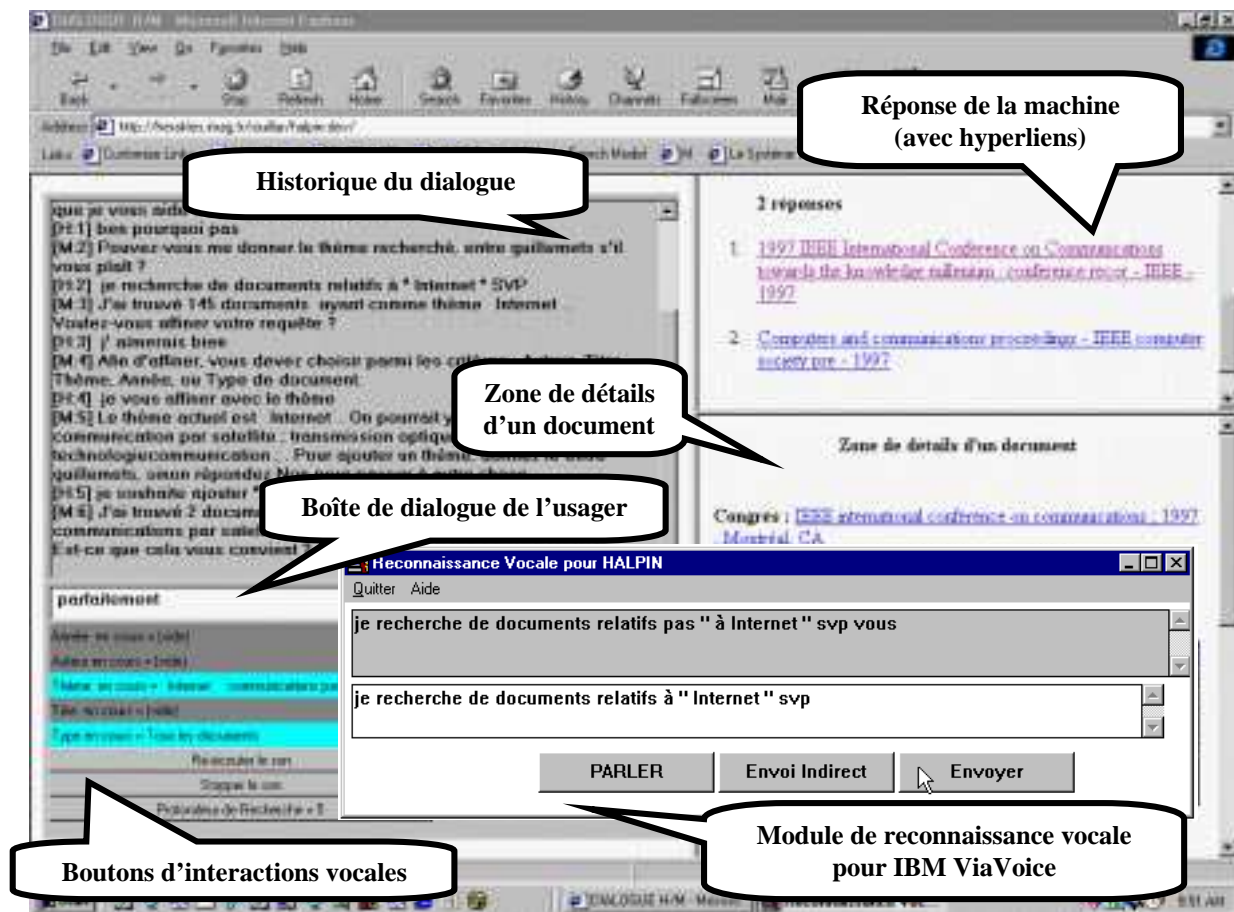


Figure 30 : L'interface du système Halpin-Documentaire pour un DHM multimodal sur le Web

Elle est divisée en trois fenêtres (dites *frames*) : celle de gauche comporte l'historique du dialogue qui permet de suivre son évolution sous forme écrite, 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, et les boutons relatifs à l'interface vocale, avec lesquels on peut arrêter ou réécouter le dernier message vocal.

A droite se trouvent les deux zones de présentation des résultats : la liste des documents d'une part, et, en dessous, le détail d'un document, d'autre part.

Au début de l'interaction, les visiteurs qui découvrent le système Halpin-Documentaire peuvent opter pour une visite guidée, dans laquelle ils se laissent diriger sur des pages de présentation hypertextuelles :

1. Présentation générale du système Halpin
2. Recherche libre ou recherche guidée
3. Objectif de votre recherche documentaire
4. Comment dialoguer avec le système ?

Ces pages contiennent également des messages sonores d'explication. L'utilisateur peut ensuite reprendre le cours normal du dialogue avec la machine. Il peut cependant demander de l'aide à

l'ordinateur (« au secours », « à l'aide », « s.o.s. », « help ») à tout moment. Ce dernier ouvre alors une fenêtre d'aide, dédiée aux pages d'explications (cf. en annexe 6 les différentes aides : sur la tâche, sur la langue, sur la parole, sur l'interface graphique).

Le système Halpin utilise donc, un processus multimodal non seulement en entrée, mais également en sortie : les réponses aux questions de l'utilisateur sont fournies de manières textuelle et hypertextuelle d'une part, et vocale d'autre part, si l'usager le désire. Il peut demander à désactiver cette option s'il ne possède pas de matériel audio ou s'il veut réduire le temps d'attente entre chaque intervention. C'est environ deux fois plus rapide sans synthèse vocale.

En résumé :

Le système Halpin-Documentaire résulte de la mise en pratique de notre modèle de DHM basé sur la reconnaissance de concepts. Il s'applique au domaine de la recherche d'information documentaire. Nous avons utilisé la base de données de l'INRIA, spécialisée dans l'informatique et l'automatique, et accessible par Internet.

Ce système multimodal permet à un usager du Web de dialoguer en temps réel, avec une machine, en langue naturelle française, afin d'effectuer des recherches bibliographiques selon des critères précis ou approximatifs.

Il met en œuvre des composantes de reconnaissance automatique de la parole, de dialogue personne-système, de recherche d'information, et de synthèse vocale. Une gestion des événements permet d'interpréter les interventions de l'utilisateur, et d'adapter la stratégie de dialogue en fonction de plusieurs paramètres.

Le chapitre suivant présente les résultats relatifs à notre système. Halpin-Documentaire est évalué, notamment par le biais de critiques obtenues auprès des utilisateurs, et par mesure des performances (comme les temps de réponses). Puis les limites du système sont discutées.

7 Résultats et évaluations du système Halpin-Documentaire

Dans ce chapitre, nous présentons des éléments permettant de critiquer et d'évaluer le système Halpin-Documentaire. Des exemples de dialogues sont présentés. L'utilité du dialogue en LN est montrée, en comparant une interrogation classique de la base de données de l'INRIA grâce à un formulaire traditionnel du Web, et une interrogation de cette même base de données par l'intermédiaire du système Halpin-Documentaire. Les appréciations et les critiques recueillies directement auprès des utilisateurs du système sont également présentées dans ce chapitre. Des résultats d'expériences concernant l'intégration, dans Halpin, d'une vue en œil de poisson, censée améliorer la présentation d'une grande masse d'informations, sont exposés et commentés. Enfin, l'évaluation du temps de réponse de la machine est donnée et les limites du système actuel sont discutées.

7.1 Présentation des résultats

La campagne d'informations menée pour le système Halpin-Documentaire, par courrier électronique et dans les news, afin que les utilisateurs testent le système, nous a donné moins de connexions que celle du système Halpin-Recueil.

Les informations relatives au trafic sur notre réseau nous indiquent que beaucoup d'usagers se sont connectés, mais nous avons dû éliminer un grand nombre de fichiers recueillis.

Il y a plusieurs raisons à cela :

- Beaucoup d'utilisateurs ont été gênés par des problèmes de *cache* et de *proxy*. (cf. note 37). Ils ont très vite abandonné l'expérience.
- D'autres n'ont pas compris qu'il s'agissait d'une recherche dans le domaine de « l'informatique et de l'automatique », et ont posé des requêtes hors contexte : « 14:19:16:H0=est-ce qu'il fera beau demain ? », « 10:18:56:H0=j'aimerais voir des images de bateaux », etc.

- Enfin, certaines personnes qui avaient déjà participé au recueil du premier corpus ont cru que Halpin-Documentaire était une version mise à jour d'Halpin-Recueil (et non pas une autre expérience de réelle recherche documentaire), et ne se sont pas reconnectées.

Cela explique que nous n'ayons obtenu que 188 dialogues exploitables. Ce qui est cependant nettement plus que ce que les procédures de type Magicien d'Oz obtiennent.

Les premiers résultats que nous exposons ci-après sont d'ordre qualitatif plus que quantitatif. Ils montrent que les utilisateurs coopèrent et dialoguent volontiers avec la machine. Cela est dû notamment au module de synthèse vocale intégré, qui semble être un fort incitateur de maintien du dialogue. L'exemple ci-dessous présentant un dialogue réel, recueilli avec le système Halpin-Documentaire, montre comment un usager du Web interagit avec la machine, en LN, pour obtenir les références de documents relatifs aux « réseaux ».

129.88.41.73 SUR Mozilla/4.0 (compatible; MSIE 4.01; Windows NT) LE 14:5:99
 15:53:32:M0=Que souhaitez-vous faire ? 1 : Retrouver un document que vous connaissez, 2 : Rechercher un ensemble de documents, ou 3 : Découvrir ce site ?
 15:53:32:H0=2
 15:53:44:M1=Rechercher un ensemble de documents. D'accord. Souhaitez-vous que je vous aide à formuler votre requête ?
 15:53:53:H1=volontiers
 15:54:03:M2=Pouvez-vous me donner le thème recherché, entre guillemets s'il vous plaît ?
 15:54:26:H2=essaye avec "réseau" pour voir
 15:54:57:M3=J'ai trouvé 2954 documents ayant comme thème réseau. Voulez-vous affiner votre requête ?
 15:55:44:H3=avec la date c'est possible ?
 15:55:56:M4=Vous pouvez soit me donner une année, soit effacer l'année en cours.
 15:56:17:H4=99
 15:56:55:M5=J'ai trouvé 13 documents de 1999 ayant comme thème réseau. Voulez-vous affiner votre requête ?
 15:58:14:H5=non
 15:58:18:M6=Dites moi si ces résultats vous satisfont.
 15:58:42:H6=oui, ca a l'air bon
 15:58:51:M7=Très bien, je suis heureux d'avoir pu vous aider dans votre recherche.
 16:00:00:H7=peux-tu écrire plus gros ?
 16:00:09:M8=Okay j'augmente la taille de la police.
 16:00:34:H8=relance la requête maintenant
 16:00:52:M9=J'ai trouvé 13 documents de 1999 ayant comme thème réseau. Voulez-vous affiner votre requête ?
 16:01:13:H9=oui, sur le theme
 16:01:44:M10=Le thème actuel est réseau. On pourrait y ajouter par exemple : langage description matériel ; conception interface ; génie logiciel ; synthèse logique ; multimédia. Pour ajouter un thème, donnez le entre guillemets, sinon répondez Non pour passer à autre chose.
 16:02:38:H10=rajoute "conception interface" alors
 16:02:56:M11=Je n'ai pas trouvé de document de 1999 ayant comme thème réseau conception interface. Voulez-vous modifier votre requête ?

16:03:17:H11=annule la derniere operation plutot
16:03:30:M12=J'ai trouvé 13 documents de 1999 ayant comme thème réseau . Voulez-vous affiner votre requête ?
16:04:14:H12=bof
16:04:26:M13=Dites moi si ces résultats vous satisfont.
16:04:49:H13=ouais
16:04:59:M14=Très bien, je suis heureux d'avoir pu vous aider dans votre recherche.

Capture 19 : Exemple N°1 d'un dialogue tiré du corpus Halpin-Documentaire

Dans cet exemple (Capture 19), on observe bien entendu des tours de paroles classiques à propos de la recherche d'informations, mais on en note également d'autres, concernant l'interface, comme l'intervention *H7=peux-tu écrire plus gros ?*, qui fait réagir la machine avec *M8=Okay j'augmente la taille de la police*.

Par ailleurs, on a ici un cas typique dans lequel le système Halpin aide l'utilisateur de manière coopérative, en tenant compte de son but. On passe donc de 2954 réponses à 13 réponses ... pertinentes. Pour cela, il lui propose des termes proches du thème initialement demandé, se focalise sur l'année voulue (il comprend que 99 signifie 1999) et s'inquiète 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, de donner une liste de mots-clefs proches du thème recherché, 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 savent pas faire.

Une des limites du système est déjà visible dans cet exemple : c'est le fait que l'utilisateur doive préciser le thème de sa recherche entre guillemets. Cette contrainte a été imposée pour détecter une partie de l'énoncé que l'on ne veut pas envoyer en phase de lemmatisation. Sans cela, une phrase comme « je cherche des documents sur les nouvelles interfaces » serait transformée en « je chercher de document sur le nouveau interface », et le thème précis demandé par l'utilisateur serait perdu.

7.2 Évaluation du système

L'évaluation d'un système de dialogue n'est pas triviale. Contrairement à d'autres disciplines, où un résultat donné est facilement appréciable (un taux de reconnaissance de la parole par exemple), il est difficilement envisageable, dans le domaine du DHM, de comparer un système à un autre. D'une part, parce que chaque système est spécialisé dans un type de tâche particulier (renseignements météorologiques, touristiques, documentaires, etc.), et d'autre part, parce que les systèmes de DHM réellement effectifs, et potentiellement comparables ne sont pas nombreux.

« Cela ramène le problème de l'évaluation à la détermination de critères permettant de décider ce qui est une bonne réponse, ou même ce qui est un bon dialogue. Cette question a priori très simple ne peut malheureusement pas recevoir de réponse simple et générale ! » [VILNAT 97].

Nous avons pu, cependant, montrer que, *ceteris paribus*⁴⁶, 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 premiers résultats montrent qu'un tel système permet de répondre aux attentes de certains usagers (les novices en particulier) grâce à un dialogue coopératif et finalisé, et d'effectuer une recherche documentaire satisfaisante là où, parfois, les requêtes avec l'interface initiale (formulaire traditionnel du Web) échouaient.

7.2.1 Exemples de DHM avec le système Halpin-Documentaire

Les exemples qui suivent permettent d'observer les possibilités offertes par le dialogue en langue naturelle dans le système Halpin-Documentaire. Comparé à un formulaire classique dans lequel on doit remplir des boîtes, cocher des cases, et où chaque requête est indépendante de la précédente, un système de DHM comme le nôtre apporte une aide supplémentaire à l'utilisateur.

Nous avons déjà vu dans l'exemple N°1 précédent (Capture 19), que l'utilisateur avait dialogué avec la machine à propos de l'interface, pour obtenir une police d'écriture plus grosse. Il avait pour cela demandé à la machine « *H7=peux-tu écrire plus gros ?* ». Nous voyons ci-après, avec la Capture 20, que l'usager arrive au même résultat, non pas avec une phrase interrogative donnant l'ordre à exécuter, mais plutôt avec une déclarative, exprimant un besoin, que l'ordinateur doit interpréter : « *H7=je ne vois rien a l'écran* ». La machine en déduit le problème à résoudre et agit sur l'interface graphique, pour servir au mieux l'utilisateur.

```
194.133.108.252 (net2.klee.fr) SUR Mozilla/4.0 (compatible; MSIE 5.0; Windows NT; DigExt) LE
5:8:99
(...)
11:55:51:M5=Vous ne voulez pas choisir de critère ?
11:57:03:H5=si, si, l'annee
11:57:14:M6=Vous pouvez soit me donner une année, soit effacer l'année en cours.
11:57:30:H6=cherche avec 95
11:57:48:M7=J'ai trouvé 34 documents de 1995 ayant comme thème intelligence artificielle .
Voulez-vous affiner votre requête ?
11:58:25:H7=je ne vois rien a l'écran
```

⁴⁶ c'est-à-dire avec une base de données identique.

11:58:35:M8=Okay j'augmente la taille de la police.
11:58:47:H8=et coupe le son
11:59:06:M9=Okay je coupe le son.
(...)
12:11:13:H14=aide
12:11:44:M15=Okay voici une fenêtre d'aide. Souhaitez-vous que je vous aide à formuler votre requête ?
12:17:42:H15=je veux une thèse de Boole
12:17:55:M16=Répondez-moi sans ambiguïté. Voulez-vous que je vous aide ou pas ?
12:18:06:H16=non
12:18:12:M17=Très bien, je vous laisse faire. Appelez à l'aide si vous êtes en difficulté.
12:18:22:H17=je veux une thèse de Boole
12:18:30:M18=Est-ce que Boole est bien l'auteur du document.
12:18:37:H18=oui
12:18:42:M19=Je n'ai pas trouvé de thèse avec comme auteur Boole.
(...)
12:19:29:M21=Vous souhaitez modifier votre requête. Selon quel critère je vous prie ?
12:19:38:H21=le type
12:19:51:M22=Vous pouvez choisir : ouvrages, congrès, thèses, rapports, ou tous les documents.
Ou dire Non pour passer à autre chose.
12:20:02:H22=tous
12:20:12:M23=J'ai trouvé 3 documents avec comme auteur Boole .
12:20:21:H23=efface Boole
12:20:31:M24=Je n'ai pas d'informations suffisantes pour lancer une requête.
12:20:48:H24=je veux un truc de Boole
12:20:57:M25=Est-ce que Boole est bien l'auteur du document.
12:21:02:H25=non
12:21:10:M26=J'ai trouvé 100 documents ayant comme thème Boole .
12:21:28:H26=je ne veux que les plus récents
12:21:39:M27=J'ai trouvé 8 documents récents ayant comme thème Boole .
(...)
12:22:26:H28=ok, imprime
12:22:37:M29=J'ai bien compris le mot imprimer, mais je ne peux pas effectuer l'opération correspondante.
(...)
12:23:42:H29=l'auteur est Caelin
12:23:53:M30=Je n'ai pas trouvé de document récent avec comme auteur Caelin ayant comme thème Boole .
(...)
12:24:32:M32=Pouvez-vous me donner le thème recherché, entre guillemets s'il vous plaît ?
12:24:40:H32=non
12:24:48:M33=Faites un choix parmi les critères : Auteur, Titre, Thème, Année, ou Type de document.
12:24:56:H33=auteur
12:25:11:M34=Vous pouvez soit m'indiquer le nom de l'auteur, avec une majuscule, soit effacer l'auteur en cours. Ou bien voulez-vous que je cherche les auteurs proches de Caelin ?
12:25:27:H34=ah oui, volontiers
12:25:48:M35=Les auteurs proches de Caelin sont caelen, jean ; caelli, terry. Donnez un nouveau nom ou dites Non pour passer à autre chose.
12:26:06:H35=bon essaye avec Caelen alors

12:26:20:M36=Je n'ai pas trouvé de document récent avec comme auteur Caelen ayant comme thème Boole. Voulez-vous modifier votre requête ?
12:26:42:H36=oui, efface le theme et la date
12:26:57:M37=J'ai trouvé 2 documents avec comme auteur Caelen. Consultez-les sur la page de droite. Est-ce que cela vous convient ?
12:27:11:H37=oui
12:27:29:M38=Très bien, je suis heureux d'avoir pu vous aider dans votre recherche.

Capture 20 : Exemple N°2 d'un dialogue tiré du corpus Halpin-Documentaire

Nous voyons qu'en H17 (je veux une thèse de Boole) et en H24 (je veux un truc de Boole), le système détecte une ambiguïté et demande si l'utilisateur souhaite travailler sur le critère thème ou sur le critère auteur. Selon la réponse, les résultats présentés sont différents.

On note que l'utilisateur a le loisir de nommer un critère explicitement, comme en H36 (*oui, efface le theme et la date*) – l'accent manquant sur le mot *thème* étant corrigé de manière automatique par le lemmatiseur -, ou bien de se référer à la valeur de ce critère dans le dialogue, comme en H23 (*efface Boole*) ; Boole mis pour auteur.

En H26 et M27, l'utilisateur utilise un élément langagier relativement flou (*je ne veux que les plus recents*) – là encore, la mauvaise accentuation est tolérée par la machine –, pour affiner sa requête. L'ordinateur interroge alors la base de données en utilisant une troncature (année=199*) ; l'étoile pouvant être remplacée par n'importe quel chiffre. Cela revient à rechercher tous les documents de la décennie en cours.

L'exemple suivant (N°3, Capture 21) est encore relatif au critère « date », et montre que la machine analyse les données fournies par l'utilisateur avant de les utiliser pour une requête. Ainsi, en M8, une année non valide est détectée, et une correction est demandée.

129.88.32.156 (jmenpet.imag.fr) SUR Mozilla/4.5 [en] (WinNT; I) LE 26:3:99
(...)
18:44:20:M7=Vous pouvez soit me donner une année, soit effacer l'année en cours.
18:44:35:H7=2001
18:44:46:M8=Vous avez dû vous tromper car 2001 n'est pas une année valide. Recommencez s'il vous plaît.
18:45:09:H8=1901
18:45:21:M9=Je n'ai pas trouvé de document de 1901 avec comme auteur Dubois. Voulez-vous modifier votre requête ?
18:45:36:H9=gros nul, cherche encore
18:45:49:M10=Je n'ai pas trouvé de document de 1901 avec comme auteur Dubois. Voulez-vous modifier votre requête ?
18:46:03:H10=je veux tous recommencer
18:46:18:M11=D'accord, tous les critères sont vides à présent. Faites un choix parmi les critères : Auteur, Titre, Thème, Année, ou Type de document.

Capture 21 : Exemple N°3 d'un dialogue tiré du corpus Halpin-Documentaire

Un autre élément important que l'on peut remarquer dans ce dialogue, réside dans le fait que la machine coopère grâce au langage, et propose de manière spontanée des informations qu'elle juge utiles pour atteindre l'objectif donné par l'utilisateur.

La Figure 31 montre l'interface traditionnelle proposée par l'INRIA pour interroger sa base de données du Web. Dans cet exemple, l'utilisateur faisant une recherche sur « Caelin » n'obtient aucune réponse, et l'interaction s'arrête là !

The screenshot shows a search interface with the following elements:

- Top left: *Cliquez sur le texte souligné pour visualiser les références complètes*
- Search input: **Vos termes** with the text "caelin" entered.
- Instruction: *Indiquez vos termes de recherche en les séparant par un espace*
- Document type selection: **Type de document** with checkboxes for:
 - Tous
 - Ouvrage
 - Congrès
 - Thèse mémoire
 - Rapports
 - Normes
- Instruction: *Vous pouvez sélectionner un ou plusieurs types de document*
- Buttons: "Nouvelle requête" and "Rechercher"
- Bottom instruction: *Cliquer sur **Rechercher** listera les réponses dans la fenêtre de gauche*
- Result area: **Aucune réponse**

Figure 31 : Interaction traditionnelle avec le formulaire de recherche de l'INRIA

Avec Halpin-Documentaire, le dialogue va plus loin, et la machine informe l'utilisateur (en M35) que *Les auteurs proches de Caelin sont caelen, jean ; caelli, terry*. Cela permet à l'utilisateur de formuler une nouvelle requête, et d'atteindre son but.

Dans la Capture 22, l'utilisateur donne le nom d'un auteur, obtient des réponses, puis décide de nommer un second auteur. En utilisant le langage naturel, il explique que ce deuxième nom est celui d'un co-auteur. La machine interprète correctement cette nouvelle information et prépare la requête adéquate (Coulon AND Kayser). L'ordinateur prend à sa charge la formulation de la requête, qui comporte un élément logique booléen, mais cela est totalement transparent pour l'utilisateur. De plus, grâce à l'historique du dialogue, on peut comparer les résultats de la première et de la seconde recherche, ce qui réduit l'effort mémoriel à fournir durant la recherche.

195.36.128.210 (Lyon-1-210.club-internet.fr) LE 2:6:99
 19:43:39:M0=Que souhaitez-vous faire ? 1 : Retrouver un document que vous connaissez, 2 : Rechercher un ensemble de documents, ou 3 : Découvrir ce site ?
 19:43:39:H0=1
 19:43:57:M1=Retrouver un document que vous connaissez. D'accord. Souhaitez-vous que je vous aide à formuler votre requête ?
 19:44:18:H1=ok
 19:44:27:M2=Connaissez-vous le nom de l'auteur du document que vous recherchez ?
 19:45:53:H2=Coulon je crois bien
 19:46:06:M3=J'ai trouvé 17 documents avec comme auteur Coulon. Voulez-vous affiner votre requête ?

19:46:38:H3=son co-auteur est Kayser
19:46:53:M4=J'ai trouvé 10 documents avec comme auteur Coulon Kayser. Consultez-les sur la page de droite. Voulez-vous affiner votre requête ?
(...)

Capture 22 : Exemple N°4 d'un dialogue tiré du corpus Halpin-Documentaire

D'autres dialogues du corpus mettent en évidence des situations coopératives où la machine intervient de manière à affiner ou élargir la recherche, afin de s'orienter vers des solutions auxquelles l'utilisateur n'aurait pas pensé au premier abord.

Ces exemples témoignent de l'utilité de notre système. Cependant, nous avons voulu connaître l'avis des principaux intéressés, comme l'explique la partie suivante.

7.2.2 Critique du système Halpin-Documentaire par les utilisateurs

Les usagers qui se sont connectés au système Halpin avaient la possibilité de répondre à un questionnaire d'évaluation facultatif, en ligne (cf. annexe 10). Nous présentons ci-dessous quelques uns des résultats obtenus, bien que l'échantillon des personnes ayant répondu à ce questionnaire de satisfaction soit assez faible (environ 30 sujets), par rapport au nombre de connexions effectivement enregistrées (188 fichiers)

La majorité des personnes ayant répondu sont des hommes (78,95 %). La plupart sont des informaticiens (57,89 %). Les bibliothécaires sont peu nombreux (11,76 %). Notons que les salariés sont représentés à 42,11 %, les enseignants et/ou chercheurs à 31,58 %, et les étudiants à 26,31 %.

Globalement, « *l'interaction dialoguée avec le système HALPIN* » est considérée, par les personnes interrogées, comme bonne à 55,56 %, mauvaise à 33,33 % et très mauvaise à 11,11 %.

A la question « *Le dialogue avec la machine vous a-t-il orienté dans votre recherche ?* », 23,53 % ont répondu « pas assez », 47,06 % « suffisamment », et 29,41 % « un peu ».

Les utilisateurs déclarent à 58,82 % que le dialogue leur a permis de trouver ce qu'ils cherchaient (*en partie* : 23,53 % ; *un peu* : 17,65 % ; *totalemment* : 17,65 %) contre 41,18 % n'ayant *pas du tout* retrouvé l'information recherchée.

Le dialogue est perçu comme robuste pour la moitié des personnes interrogées, et 56,25 % ont apprécié l'aide apportée par la machine lors des interactions dialoguées.

La synthèse vocale utilisée est considérée comme « *très bonne* » pour 21,05 % des utilisateurs, « *bonne* » pour 47,37 % et « *mauvaise* » pour 10,53 % (21,05 % de sans opinion). Elle leur semble très intelligible (42,11 %) et moyennement naturelle (36,84 %).

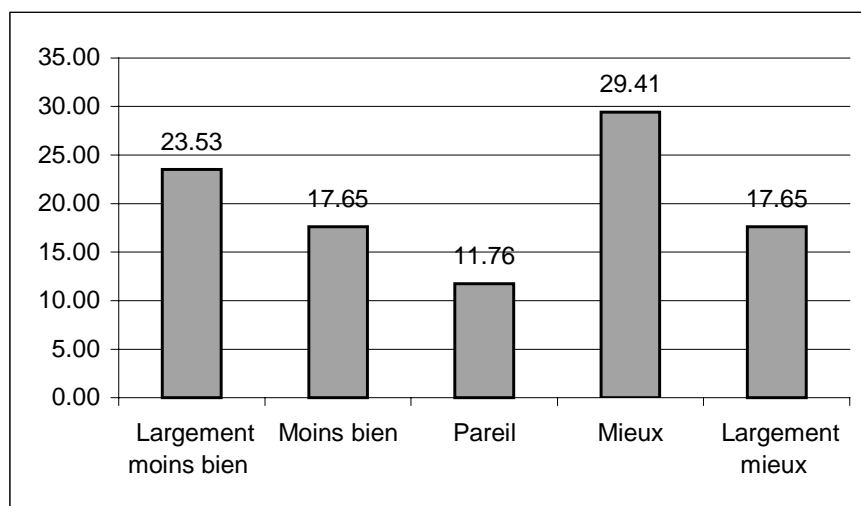


Figure 32 : Utilité du système Halpin par rapport à un formulaire traditionnel du Web

Les avis à propos de l'utilité du système, comparé à un formulaire traditionnel du Web, sont présentés sur la Figure 32 ci-dessus.

Dans l'ensemble, 41 % des personnes trouvent ce système moins utile qu'une interaction traditionnelle (23,63 % *largement moins bien* et 17,65 % *moins bien*), 12 % le trouvent similaire et 47 % le trouvent meilleur (*mieux* à 29,41 % et *largement mieux* à 17,65 %).

Le temps de réponse du système, quant à lui, est jugé comme étant *rapide* à 15,79 %, *acceptable* à 36,84 % et *trop long* à 47,37 %.

Nous avons vu que la plupart des personnes ayant répondu à ce questionnaire sont des informaticiens. Même si ces résultats sont conformes à nos hypothèses, les critiques des utilisateurs peuvent être nuancées par le fait que, précisément, elles sont formulées, en grande partie, par des experts du domaine informatique. Il est concevable qu'ils préfèrent une interaction traditionnelle, avec laquelle ils ont l'habitude de travailler, plutôt qu'une interaction dialoguée comme celle que nous avons présentée. Une étude auprès de véritables novices (des personnes ne connaissant pas Internet et le monde de l'informatique), donnerait probablement des résultats différents.

7.2.3 Évaluation du temps de réponse de la machine

Un élément significatif lors de l'évaluation d'un système de DHM réside dans temps de réponse de la machine, à chaque interaction. Cela est d'autant plus important pour notre système qui fonctionne sur le Web.

Comme nous l'avons vu, dans Halpin-Documentaire, plusieurs étapes sont requises pour que la machine réponde à l'humain. *A priori*, celles qui nécessitent le plus de temps sont l'analyse morphologique (suivie de la reconnaissance de concepts), l'éventuel accès à la base de données de l'INRIA, et le calcul de la synthèse vocale.

Nous avons voulu savoir dans quelle proportion chacun de ces modules intervenait dans le calcul de la réponse, et combien de temps prenait, en moyenne, l'élaboration de cette réponse globale. Les résultats obtenus ne tiennent pas compte de la reconnaissance de la parole, car elle s'utilise indépendamment du navigateur (dans une autre fenêtre) : on ne peut pas capturer à distance les événements d'une fenêtre cliente.

Temps	Global (sec)	Xérox (sec)	% Xérox	Inria (sec)	% Inria	Élan (sec)	% Élan
Moyenne	20.17	4.57	25.23	6.17	25.25	9.17	47.79
Écart type	6.12	0.67		3.40		4.94	
Minimum	14.00	4.00	13.79	1.00	6.25	7.00	30.43
Maximum	32.00	6.00	37.50	11.00	47.83	19.00	65.52

Tableau 17 : Analyse statistique du temps nécessaire pour une interaction Homme-Machine

Comme l'indique le Tableau 17 ci-dessus, le temps moyen requis pour une interaction dialoguée globale via le Web est de 20.17 secondes, entre le moment où l'utilisateur valide son énoncé et celui où il entend la réponse de la machine.

La Figure 33 montre comment ce temps se répartit entre les différents modules étudiés, lorsqu'il y a accès à la base de données. L'analyse morphologique demande 26 % du temps. La reconnaissance de concept est très rapide (moins d'une demi-seconde) et n'est pas mentionnée sur le graphique. L'accès à la bibliothèque électronique compte également pour 26 % du temps global, et la synthèse vocale équivaut à presque la moitié du temps complet (48%).

Lorsqu'il n'y a pas d'accès au serveur de l'INRIA, une interaction complète prend en moyenne 12.36 secondes. Cela se divise en 45 % pour l'accès à Xérox et 55 % pour la préparation de la synthèse vocale.

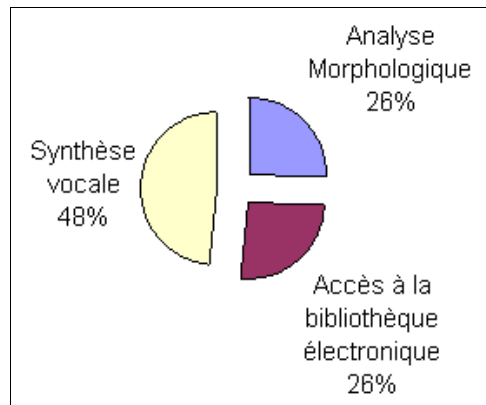


Figure 33 : Temps relatif nécessaire pour fournir une interaction dialoguée sur le Web

Nous voyons donc que l'amélioration des temps de réponse passe par une réduction du temps nécessaire pour le calcul de la synthèse vocale. Dans cette perspective, l'arrivée de nouvelles technologies permettant de faire synthétiser la voix directement par le navigateur client, dans une applet Java, est à étudier (cf. Java Speech API⁴⁷).

7.2.4 Intégration d'une vue en œil de poisson

On voit se développer depuis quelques années des travaux tendant à aider l'utilisateur lorsqu'il se trouve confronté aux problèmes de surcharge cognitive et de désorientation sur le Web. [BOURDONCLE 97] propose ainsi une visualisation graphique (cf. Figure 34 ci-après) des résultats obtenus lors d'une recherche d'informations sur le moteur Altavista⁴⁸.

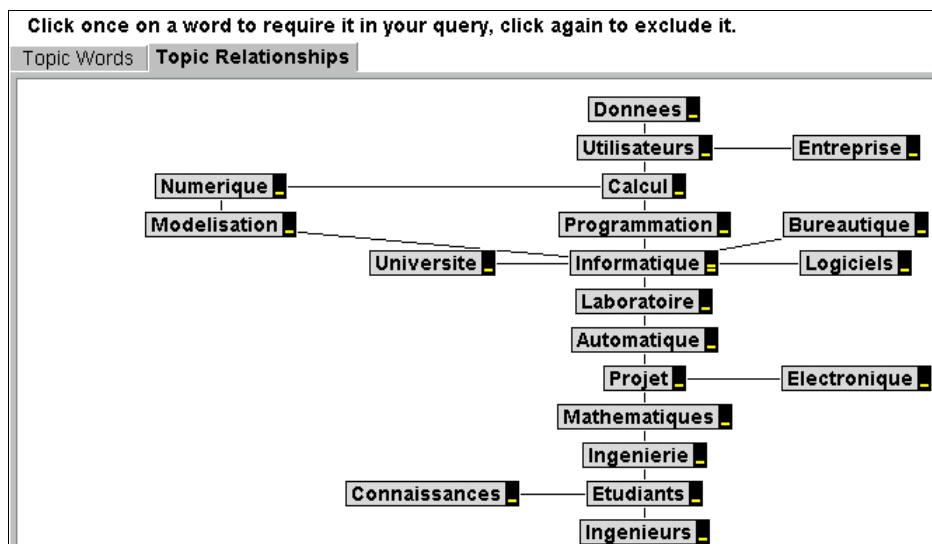


Figure 34 : Représentation graphique de Livetopics pour le moteur de recherche Altavista

⁴⁷ <http://java.sun.com/products/java-media/speech/>

⁴⁸ <http://www.altavista.com/>

Des travaux similaires, sur la multimodalité, ont été menés, avec le système Halpin-Documentaire, dans le cadre du projet RICOM⁴⁹ (Recherche d'Information par interaction COopérative Multimodale) du CNRS et de l'IMAG. Une vue en œil de poisson (dite fish-eye-view ou FEV) a notamment été intégrée dans la zone de présentation des résultats, comme on peut le voir sur la Figure 35 ci-après. Dans cette étude, nous avons essayé de déterminer dans quelle mesure et pour quelles tâches les utilisateurs préfèrent un dialogue en langue naturelle avec une machine, plutôt qu'une navigation libre.

Concernant la navigation, nous avons collaboré avec [VERNIER et NIGAY 97] pour choisir la meilleure représentation de grande quantité d'information, adaptée à notre interface.

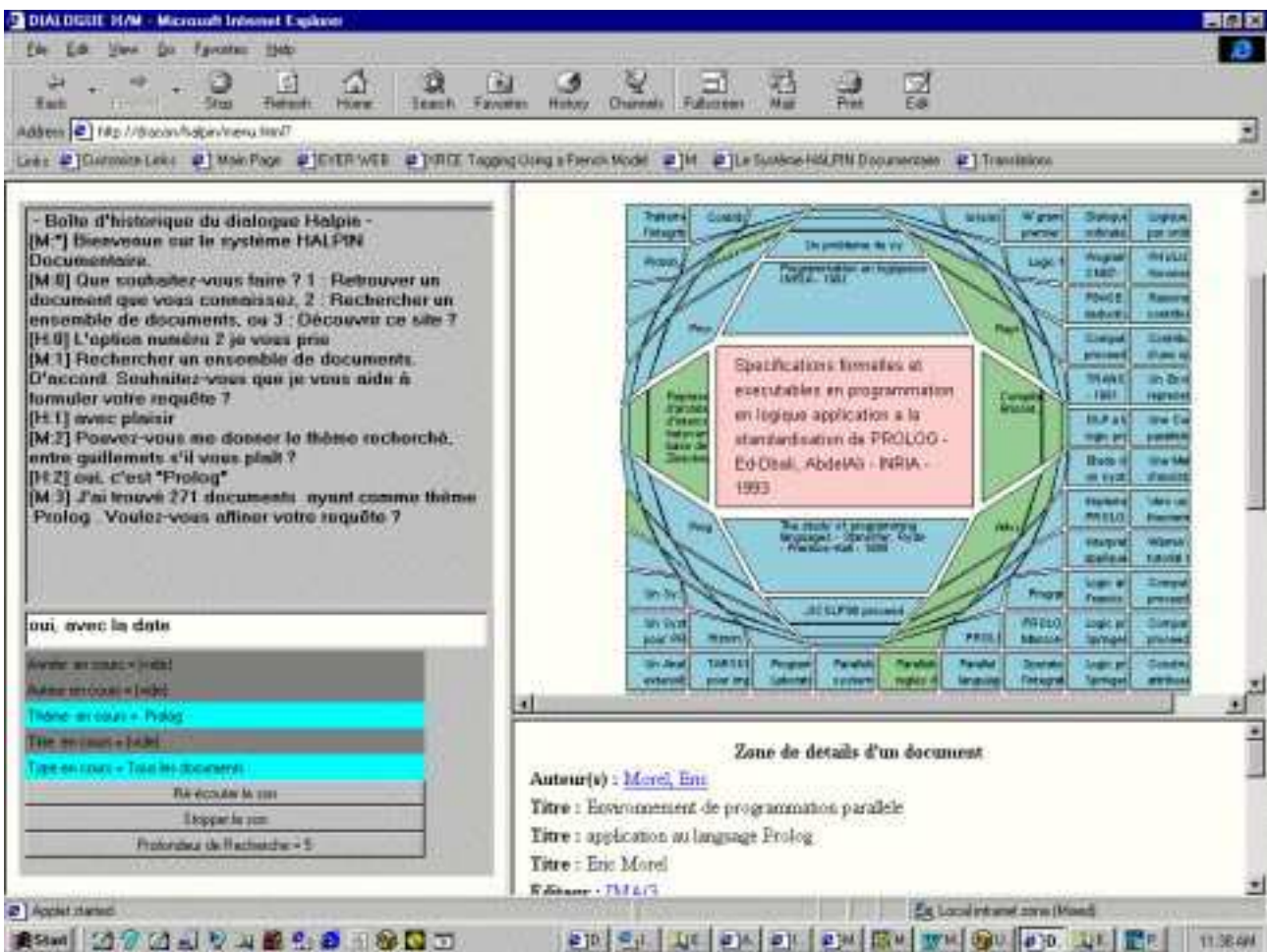


Figure 35 : Interface du système Halpin-Documentaire, intégrant une vue en œil de poisson, dite « en colline »

Nous avons décidé d'utiliser, parmi les multiples choix possibles de FEV (ovale, cylindre, sphère, carré, etc.), une vue dite « en colline » car elle donne une bonne idée du nombre total de réponses tout en permettant de focaliser assez bien sur quelques éléments (ceux au cœur de la boule). Lorsque l'utilisateur pointe à la souris un rectangle de la FEV, celui-ci devient le point culminant de la

⁴⁹ <http://iihm.imag.fr/nigay/RICOM/>

colline. On peut également déplacer la souris tout en maintenant le bouton gauche appuyé, et la surface ainsi focalisée se déplace de manière fluide.

7.2.4.1 Dialogue versus Navigation

L'objectif des expériences menées était d'évaluer l'aide apportée (ou pas) par le dialogue et la navigation (FEV). Notre hypothèse était que la classification des documents retournés par un système de recherche d'information documentaire, devait permettre d'obtenir une vision plus claire des réponses et une amélioration des performances liées à la réduction de la charge cognitive de l'utilisateur, en jouant, par exemple, sur l'ordre de présentation des documents.

Nous avons proposé dans ces expériences 5 exercices de recherche documentaire, sans limite de temps. Un expérimentateur donnait en début de séance la liste écrite des exercices à faire et quelques informations sur le fonctionnement du système, puis restait dans la salle avec le sujet, mais ne devait plus intervenir, sauf en cas de problème technique. Les bonnes réponses étaient connues des expérimentateurs⁵⁰, et leurs tâches consistaient, *a posteriori*, à déterminer comment et pourquoi tel mécanisme était préféré à tel autre, à tel moment. Autrement dit, il s'agissait de comprendre dans quelles circonstances l'utilisateur utilise plus volontiers un dialogue qu'une navigation, ou l'inverse. Les répliques avec la machine et les manipulations à la souris ont été capturées et analysées. En fin de séance, les utilisateurs devaient également répondre à un questionnaire d'évaluation et de critique des interfaces utilisées. Les personnes ayant répondu sont presque toutes étudiantes en DEUG MASS ; les 3/4 sont des hommes et la plupart - 83,33 % - disent posséder de bonnes connaissances en informatique. Voici les résultats obtenus avec les 10 personnes volontaires.

⇒ 1^{er} exercice

« Rechercher le titre d'un *rapport* (type de document) écrit par *Pierrel* en 1992 ». Le temps moyen (chronométré manuellement) pour effectuer cet exercice fut de 6,23 min. avec un minimum à 4,5 min. et un maximum à 7,27 (ce sujet à d'ailleurs abandonné l'exercice). Deux sujets sur dix n'ont pas du tout utilisé la FEV pour retrouver le bon document. Tous les autres ont d'abord dialogué puis ont navigué sur la boule. Un seul a intercalé deux phases de dialogues avec deux phases de navigation. Cet exercice s'est soldé par 8 succès et 2 échecs (abandon).

⇒ 2^{ème} exercice

« Recherchez un document ayant pour thème *parole* et édité par l'INRIA entre 1990 et 1993 ». Le temps moyen pour effectuer cet exercice fut de 9,17 min. avec un minimum à 4 min. et un maximum à 16,39. Notons que l'information concernant l'éditeur était visible dans le détail d'un document, mais qu'il n'était pas possible de soumettre une requête sur ce critère à la bibliothèque électronique de l'INRIA. Les sujets ont tous utilisé le dialogue pour affiner leurs requêtes selon le critère année. Même en se rendant compte, au fil du dialogue, qu'il n'était pas possible d'interroger la base de données de l'INRIA avec une fourchette de date (90-93), ils ont quand même persisté à dialoguer, puis ont navigué, successivement, au sein des 4 FEV (une pour chaque année). On a ainsi détecté une moyenne de 53 clics sur la boule (minimum : 28, maximum : 84) dans cet exercice.

⇒ 3^{ème} exercice

Il se décompose en deux tâches relativement proches ; avec, puis sans utilisation de la FEV, qui sont toujours à effectuer dans cet ordre :

a) « Rechercher le nombre de documents écrits en anglais ayant pour thème *dialogue* » (avec FEV).

Là encore, il n'était pas possible d'effectuer une requête prenant en compte la langue du document recherché. La FEV permet de visualiser simultanément plusieurs ouvrages et aurait pu être utilisée de manière optimale en la positionnant de manière à voir rapidement si les titres des documents semblaient anglais ou pas. Or, on constate une moyenne de 32 clics (le FEV comptait 35 facettes dans cet exercice) ; c'est à dire que les sujets ont vérifié presque chaque fiche, en focalisant le centre de la FEV sur le polygone correspondant.

b) « Rechercher le nombre de documents écrits en anglais ayant pour thème *discours* » (sans FEV).

C'était donc le même exercice mais avec une liste d'hyperliens à la place de la FEV. Le temps moyen pour réaliser cet exercice diminue sensiblement par rapport à l'exercice précédent (2,10 min. contre 4,38 min., avec un minimum à 48 sec. contre 2.32 min. précédemment). Les sujets ont expliqué qu'il était plus facile de dénombrer les documents dans la liste car on voyait 8 références simultanément à l'écran, que l'on pouvait immédiatement identifier leurs langues et que l'ascenseur permettait de faire défiler rapidement les 35 réponses pour compter le nombre de documents anglais demandés.

⁵⁰ Il s'agissait de Marie-laure Betbeder et de Fabienne Baciocchi, stagiaires de maîtrise MASS, qui ont largement contribué aux travaux que nous présentons ici.

⇒ 4^{ème} exercice

« Rechercher, dans les documents ayant pour thème la *programmation par objets*, 5 sous-thèmes (mots-clés) souvent proposés ». Il était possible de faire faire cet exercice à la machine en dialoguant avec elle, mais cela n'a pas été fait par les utilisateurs, soit parce qu'ils ne savaient pas que la machine pouvait le faire pour eux, soit parce qu'ils n'ont pas cherché à déléguer cette tâche. Ils ont donc tous procédé par navigation. On observe là encore 8 succès et 2 échecs.

⇒ 5^{ème} exercice

a) « Rechercher le plus ancien document, le plus récent et un daté de 1992, ayant pour thème *l'intelligence artificielle* (les documents présentés étant classés linéairement par ordre décroissant des dates dans la FEV) ». Quatre sujets ont dialogué avec la machine pour affiner la requête selon l'année désirée. Les autres ont cliqué entre 4 et 11 fois pour obtenir le même résultat.

b) « Rechercher le plus ancien document, le plus récent et un daté de 1983, ayant pour thème *Prolog* (les documents étant présentés et classés "en forme d'escargot" par ordre décroissant des dates dans la FEV) ». La moitié des sujets a choisi une stratégie de dialogue pour obtenir l'année voulue, tandis que l'autre moitié remplissait cette tâche en manipulant la boule, pour faire apparaître le titre, l'auteur et la date. Quatre personnes ont cliqué plus que les deux fois nécessaires pour retrouver le plus récent et le plus ancien des documents. Pour ceux-là, on note deux fois plus de clics que nécessaire avec la représentation en escargot plutôt qu'avec une représentation linéaire (le plus récent en haut à gauche et le plus ancien en bas à droite) dans la FEV. Certains ont expliqué que c'était plus délicat avec l'escargot car en manipulant la boule, au bout de quelques instants, on n'est plus en mesure de vérifier que la face du polygone visible à l'écran est bien celle qui était initialement au centre, c'est à dire celle contenant le document le plus récent. Enfin, les temps moyens constatés pour réaliser ces exercices (a) et (b) sont similaires ; respectivement 4.11 et 4.14 minutes.

7.2.4.2 Réponses au sondage concernant la FEV

Nous présentons ci-dessous quelques uns des résultats obtenus grâce au sondage effectué juste après l'expérience.

➤ Le dialogue

Le Tableau 18 donne l'opinion des utilisateurs concernant le DHM du système Halpin. Les questions correspondant aux lignes du tableau sont les suivantes :

A : « l'interaction dialoguée avec le système vous semble... »

B : « le dialogue avec la machine vous a-t-il orienté dans votre recherche ? »

C : « le dialogue vous a-t-il permis de retrouver ce que vous cherchiez ? »

A	très bonne 8,33 %	bonne 66,67 %	mauvaise 16,67 %	très mauvaise 8,33 %
B	pas assez 8 %	un peu 33 %	suffisamment 50 %	Beaucoup 8 %
C	en partie 83,33 %	un peu 8,33 %		pas du tout 8,33 %
D	très bonne 16,67 %	bonne 50 %	mauvaise 33,33 %	
E	très 25 %	assez 33,33 %	un peu 25 %	pas du tout 16,67 %
F	totallement 25 %	en partie 58,33 %	un peu 8,33 %	pas du tout 8,33 %

Tableau 18 : Opinion des usagers à propos du système utilisé

Le dialogue est perçu comme robuste pour la moitié des personnes interrogées, et 91,67 % disent avoir apprécié l'aide apportée par la machine lors des interactions dialoguées. Un tiers des personnes trouvent cependant cette interaction dialoguée moins utile qu'une interaction traditionnelle (au clavier et la souris), 8,33 % la trouvent similaire et 58,33 % la trouvent meilleure.

➤ **La navigation**

Dans le Tableau 18, la ligne D représente la satisfaction des usagers en ce qui concerne la navigation par FEV. En ligne E, les utilisateurs s'expriment à propos de son efficacité. Enfin, en ligne F, les sujets évaluent dans quelle mesure cette FEV leur a permis de retrouver ce qu'ils cherchaient.

➤ **L'utilité de la « Fish-Eye-View »**

Il apparaît que 41 % des usagers préfèrent la FEV, tandis que 51 % affectionnent davantage l'affichage traditionnel linéaire (liste d'hyperliens). Le temps de réponse du système, pour calculer et afficher la FEV dans l'applet, est jugé comme acceptable (58,33 %). La manipulation de la FEV à la souris est considérée comme difficile par 1/3 des personnes.

7.2.4.3 Conclusion sur l'intégration de la FEV dans Halpin-Documentaire

Notre interprétation des résultats obtenus nous conduit à penser que les sujets ne changent pas facilement de modalité en cours de tâche. La communication se faisant à l'initiative de la machine et en LN écrite et orale, ils ont tendance à vouloir continuer à dialoguer et ne passent à une navigation qu'en dernier recours et lorsque cela devient nécessaire. On a pu noter par exemple qu'un seul des sujets a manipulé la FEV dès son apparition à l'écran. Les autres attendaient (le temps du calcul de la synthèse vocale) que la machine leur indique oralement le nombre de documents retrouvés, avant de prendre connaissance des résultats sur la FEV. Il apparaît également que l'utilisation de la FEV n'est perçue de manière pertinente que lorsque le type de tâche à réaliser nécessite une présentation sémantiquement (pour le thème par exemple) ou chronologiquement structurée.

L'utilisation d'un outil tel que la FEV pour la représentation graphique d'une grande masse d'information n'est pas aussi « naturelle » que l'on aurait pu l'imaginer. De plus, rien ne nous laisse croire au vu de nos tests et des informations données par les sujets, que la surcharge cognitive, généralement constatée lors de navigation en environnement hypertextuel, soit réduite grâce à cet outil. Notre hypothèse de départ qui consistait à penser que l'utilisation de la FEV permettrait de réduire sensiblement le temps d'interaction est fautive. A notre avis, pour être efficace, son utilisation doit être subordonnée à une structuration hiérarchique optimale.

Par exemple, on pourrait colorier les polygones contenant des documents de la même décennie avec une couleur identique ; ou encore corréler l'intensité de la couleur (plus ou moins foncée) en fonction de la pertinence des mots-clés des documents retrouvés, comparés aux éléments de la requête. Même si l'approche de notre étude comporte quelques limites (choix des sujets, paramètres graphiques relatifs à la couleur et à la taille des polygones, par exemple), les résultats montrent que la communication en LN est ici préférée à la FEV. L'utilisation d'un DHM, en dépit des efforts de mise en œuvre qu'il demande, donne des résultats relativement satisfaisants. Les usagers négocient, affinent, et réorientent donc leurs requêtes plus aisément en dialoguant qu'en navigant.

7.3 Limites du système Halpin-Documentaire

Les limites constatées de notre système sont principalement de deux ordres : techniques et dialogiques.

7.3.1 Limites techniques

Les limites et les contraintes d'utilisation sont en fait 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 est cependant d'un niveau de performance acceptable après une période d'apprentissage du locuteur (200 phrases à prononcer).

Une autre limite réside dans le temps nécessaire au système pour formuler une réponse orale aux énoncés des utilisateurs. Nous avons vu 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 n'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, 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. Il semble que le flux des échanges entre les allocutaires soit moins ralenti par le transit des messages vocaux que par 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. 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, et par transitivité, peu naturelle, puisque moins spontanée.

7.3.2 Limites dialogiques

Nous avons déjà expliqué pourquoi nous avons décidé d'utiliser des guillemets pour indiquer les titres et les thèmes, et une majuscule pour les noms d'auteurs. Cela permet à la machine de traiter plus facilement les énoncés des usagers, mais il est vrai que c'est une limite ; qui se vérifie d'autant plus à l'oral : on ne peut guère préciser vocalement que tel mot contient une majuscule⁵¹, et le fait d'intégrer des guillemets (même si c'est techniquement possible en prononçant le mot guillemet), ôte une part de l'interaction naturelle recherchée.

Autre limite constatée : la puissance du langage naturel surpasse souvent les capacités du système opératoire, de sorte que certaines commandes sont parfaitement interprétées par le système Halpin mais impossibles à exécuter. C'est le cas avec ces quelques exemples extraits du corpus :

- « Imprime moi les 3 dernières listes de résultats » (cause : impossibilité de déclencher une impression sur une imprimante distante, sans autorisation).
- « 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 : impossibilité 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'utilisateur que sa phrase a bien été comprise, mais que l'opération correspondante ne peut pas être effectuée.

La robustesse du dialogueur est également un facteur important permettant d'évaluer un système de DHM. La plupart des dialogues du corpus sont cohérents, mais parfois, comme le montre la Capture 23 ci-dessous, l'ordinateur est encore loin des performances d'un être humain.

```
193.105.167.66 (horus.lcfr.fr) SUR Mozilla/4.0 (compatible; MSIE 4.0; Windows NT) via proxy gateway CERN-HTTPD/3.0 libwww/2.17 LE 2:4:99
11:19:08:M0=Que souhaitez-vous faire ? 1 : Retrouver un document que vous connaissez, 2 : Rechercher un ensemble de documents, ou 3 : Découvrir ce site ?
11:19:08:H0=3
11:19:23:M1=Dans la page de droite, je vous propose une visite guidée du système Halpin Documentaire. Il suffit de lire, et de cliquer sur les hyperliens.
11:21:51:H1=base de données sur le web, comment dois-je procéder ?
11:22:01:M2=Je n'arrive pas à déterminer votre but. Choisissez 1, 2 ou 3.
11:22:48:H2=2
```

⁵¹ L'aide en ligne d'IBM ViaVoice explique « Vous pouvez mettre en majuscule l'initiale d'un mot au cours de la dictée en prononçant majuscule_initiale, suivi du mot dont vous souhaitez mettre l'initiale en majuscule. »

11:23:01:M3=Rechercher un ensemble de documents. D'accord. Souhaitez-vous que je vous aide à formuler votre requête ?
11:28:18:H3=non
11:28:26:M4=Très bien, je vous laisse faire. Appelez à l'aide si vous êtes en difficulté.
11:30:42:H4=je veux des informations sur la connexion de bases de données à un site web
11:30:55:M5=Je prépare une requête ayant comme thème - base - . Nous sommes bien d'accord ?
11:32:39:H5=non
11:32:44:M6=Alors, reformulez votre phrase s'il vous plaît.
11:34:16:H6=bases de données
11:34:26:M7=Donnez-moi plus d'indications à propos de ce document.
11:38:02:H7=tu m'énerves !!!!
11:38:14:M8=Je ne comprends pas le sens de votre phrase. Reformulez je vous prie.
11:39:51:H8=le système d'exploitation windows 2000 est-il intéressant ?
11:40:00:M9=2000 n'est pas une année valide.

Capture 23 : Exemple de manque de robustesse dans le dialogue

Par ailleurs, le système Halpin-Documentaire n'est pas capable, à l'heure actuelle, de répondre à une question de l'utilisateur par une autre question, comme il est courant de le faire en conversations spontanées. [BILANGE 92] en donne un exemple complexe avec :

A : allo

B : est-ce que Eric est là ?

A : de la part de qui ?

B : il est là ?

A : oui

B : c'est sa sœur

A : un moment s'il vous plaît

Il n'est pas capable de mettre en œuvre des stratégies de réparation ou de gestion d'erreurs. Il ne sait pas interpréter ou générer des ellipses, et n'a pas la faculté d'apprendre des éléments nouveaux, au cours du dialogue, de le mémoriser et de s'en servir lors de nouvelles sessions. Il utilise son propre vocabulaire, et ne tente pas de reprendre celui utilisé par son interlocuteur. Il n'utilise pas encore les pleines capacités d'une multimodalité, dans laquelle on pourrait, par exemple, cliquer sur un lien à la souris, et prononcer simultanément une commande orale liée à l'objet ainsi désigné. Il n'exploite pas l'historique pleinement, puisqu'il ne saurait répondre correctement à une question du type « quelle était la 2^{ème} réponse déjà ? ».

Ces faiblesses résultent de contraintes et de choix que nous avons faits. Nous aurions pu nous focaliser entièrement sur la gestion du dialogue, et ne proposer qu'un système utilisable au clavier, sur un ordinateur isolé. Mais nous avons à cœur de proposer un système de DHM en langue naturelle française, réellement utilisable sur Internet, compatible avec tout type de navigateur, intégrant une synthèse vocale de qualité et autorisant une reconnaissance de la parole en entrée du système.

Une contrainte supplémentaire nous était également imposée par le fait que notre logiciel devait pouvoir être intégré dans le navigateur NESTOR, développé par le pôle « visuel » (lyonnais) du projet Orion. La Capture 24 de l'annexe 11 prouve que cette contrainte a été respectée.

Le système Halpin-Documentaire, est donc largement perfectible d'un point de vue dialogique. Le défaut que nous tenterons de corriger en priorité est relatif à la stratégie de reconnaissance de concepts adoptée. En effet, l'identification des concepts s'effectue de façon linéaire, et le dialogueur réagit de manière événementielle aux concepts détectés dans leur ordre d'apparition au sein de la phrase. Il serait intéressant de disposer d'un algorithme permettant de gérer les priorités de traitements, sémantiquement.

Ainsi, actuellement, à la question « Voulez-vous affiner votre requête ? », si l'utilisateur répond « Oui, mais d'abord, je veux revoir la liste précédente », le système va détecter un concept d'affirmation dans la phrase de son interlocuteur, et lancer le processus d'affinement d'une requête, ce qui ne va satisfaire l'utilisateur que partiellement.

En résumé :

Dans ce chapitre consacré à la présentation des résultats et à l'évaluation du système, nous avons vu que l'adaptation de la machine par rapport à son interlocuteur (niveau d'assistance, but, etc.) rend le DHM plus crédible et plus performant, et que dans certaines situations, ce type d'interaction donne de meilleurs résultats que lors d'une interaction traditionnelle, dans laquelle l'utilisateur remplit un simple formulaire sur le Web.

En suivant la logique du processus incrémental de notre projet, des évaluations du système Halpin-Documentaire ont été menées, d'une part grâce à un questionnaire que les personnes connectées remplissaient de manière volontaire, juste après l'expérience, et d'autre part en observant des données (les temps de réponse du système notamment), de manière systématique, lors des connexions des utilisateurs.

Le principal reproche de certains utilisateurs du système réside dans le temps d'attente, jugé parfois trop long. Les observations effectuées indiquent, entre autres, que la synthèse vocale calculée à la volée, représente près de la moitié du temps nécessaire à l'élaboration de la réponse de la machine.

Malgré tout, la plupart des utilisateurs trouve ce type d'interaction avec une machine intéressant, voire meilleur qu'une interaction traditionnelle.

8 Conclusion et perspectives

◆ Bilan scientifique

Dans le cadre du projet Orion, nous nous sommes intéressé à l'étude de nouvelles formes d'interactions pour la navigation et la recherche d'informations sur le Web. Cette thèse avait pour but de rendre compte des difficultés rencontrées par les utilisateurs naviguant et recherchant de l'information sur le Web, et devait permettre d'explorer des pistes dans le domaine du dialogue Homme-Machine, écrit et oral.

Un état de l'art des systèmes de DHM a permis de constater que les supports les plus employés sont le téléphone et l'ordinateur, que la plupart des applications permettent d'accéder à des bases de données explicites limitées, d'effectuer, grâce au dialogue, des actions sur ces BD, et enfin que la moitié des systèmes de DHM considère l'utilisateur de manière générique, tandis que l'autre moitié des systèmes s'adapte dynamiquement à l'usager, au cours du dialogue.

D'autre part, il a été montré que les outils actuellement disponibles pour la recherche d'information sur le Web demandent à l'utilisateur de fournir un effort cognitif pour traduire une question formulée en langage naturel, en une suite de mots-clés combinées à des opérateurs logiques booléens. Dans tous les moteurs de recherche étudiés, chaque requête que formule l'usager est indépendante de la précédente. De plus, aucun de ces systèmes ne s'adapte à son interlocuteur pour lui fournir un meilleur service.

Une enquête d'usage des outils informatiques pour la recherche d'informations, réalisée dans une médiathèque, a démontré que les utilisateurs ont des difficultés à retrouver des données pertinentes avec les moyens actuellement proposés, c'est-à-dire avec des interfaces manipulables au clavier et la souris, dans lesquelles on effectue des recherches en remplissant des boîtes dans un formulaire. Cette enquête concluait à la possibilité d'utiliser des interfaces multimodales, s'appuyant sur un dialogue Homme-Machine en langue naturelle.

Sur ces constats, nous avons décidé de réaliser un système de DHM utilisant la langue française, pour rechercher de l'information sur une partie du Web, et capable de s'adapter à son interlocuteur de manière dynamique, au cours du dialogue. Nous avons choisi d'utiliser le réseau Internet comme support d'expérimentation, d'une part pour développer un système utilisable et critiquable par le plus grand nombre de personnes, et d'autre part, afin de recueillir des corpus de DHM selon une méthodologie autre que celle du Magicien d'Oz.

Nous avons inscrit notre démarche de conception dans un processus incrémental, en spirale. On améliore ainsi l'application efficacement, au vu des critiques et des utilisations réelles des usagers. Ces derniers se connectent de toute la France, voire de l'étranger, et donnent alors une meilleure couverture géographique au projet. Des expériences reprenant cette méthodologie de recueil de corpus sur Internet, sont utilisées dans le domaine des sciences cognitives [MEIYE et al. 98] et dans celui du DHM [LEMEUNIER 99].

Nous avons modélisé le DHM en nous basant sur le modèle COR (cf. annexe 12) auquel nous avons ajouté la prise en compte d'éléments importants pour la pertinence du dialogue, comme le profil de l'utilisateur (ses compétences, ses besoins) et le but de sa recherche, afin que la machine puisse, en analysant ses énoncés, faire émerger des concepts et des actes de langage. De manière coopérative, mais sans être trop intrusive, elle réplique efficacement, justifie cette réponse, et suggère des alternatives pour la poursuite du dialogue, devant mener à l'accomplissement de la tâche.

Nous avons implémenté une application de dialogue réel entre un usager et un ordinateur, fonctionnant sur le réseau Internet, et permettant, dans un premier temps, de recueillir un corpus conséquent pour l'élaboration des modèles de l'utilisateur, de la langue, et du dialogue, puis dans un second temps, de fournir un système prenant en compte les besoins et les profils des usagers, afin de mieux les guider dans leurs recherches.

Ces travaux semblent valider notre hypothèse de base, à savoir que l'aspect sémantique est une composante forte pour un dialogue coopératif et finalisé, et qu'un modèle de DHM fondé sur une reconnaissance de concepts est viable. Il n'est pas déraisonnable de penser que notre système pourrait être porté, à moindre coût, vers d'autres domaines, en adaptant le vocabulaire de la tâche.

Une partie des concepts est fixe, et commune à tout type d'application, et une autre partie est relative à la tâche. Ainsi, grâce à cette structure de concepts, pour porter une application d'un

domaine vers un autre, la première étape consisterait à remplacer le lexique de la tâche actuelle par celui de la tâche visée. D'autres équipes, comme celle du CMU (Carnegie Mellon University) ont adopté la même démarche : « *We took a completely semantic approach to this problem. (...) Our parser is domain-limited but very robust. (...) The process is easily ported to others domains.* » [MAYFIELD et al. 95]. Il faudrait ensuite connecter le système à la base de données concernée, via Internet, réécrire le module qui permet de récupérer les informations fournies en langage HTML, et adapter celui qui gère les événements relatifs au domaine spécifié.

Pour ces travaux, 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.

◆ **Bilan technologique**

Nous avons mis au point un premier système informatique (Halpin-Recueil) afin de recueillir un corpus de DHM sur le Web. Nous avons ainsi obtenu plus d'un millier de fichiers, fournissant un corpus de dialogues que nous avons jugé suffisant. Cela nous a permis de valider notre hypothèse et d'étudier les paramètres linguistiques, dialogiques et socio-cognitifs que l'on pouvait déduire de ce corpus.

Puis, les différentes composantes nécessaires pour un DHM oral sur Internet ont été intégrées : reconnaissance vocale, compréhension du LN par identification de concepts, et synthèse vocale. Enfin, nous avons éprouvé notre modèle, grâce à une réalisation logicielle complète de DHM, avec entrées et sorties vocales via le Web : le système Halpin-Documentaire. Ce système, multiplateforme, qui permet d'interroger une base de données d'Internet, est utilisable avec n'importe quel navigateur compatible Java.

Il a permis de montrer qu'un tel hyperdialogue est apprécié par une partie des usagers, notamment les débutants. Son utilisation, basée sur la langue naturelle et l'emploi d'un dialogue oral, permet de réduire la surcharge cognitive de l'utilisateur, et dépasse, dans certains cas, les performances que l'on peut atteindre avec une interface traditionnelle.

Les utilisateurs qui n'ont pas apprécié ce système dénoncent, principalement, les temps d'attente nécessaires à l'élaboration d'une réponse orale, et quelquefois le manque de robustesse de la part de la machine, pour prétendre entretenir une conversation significative avec un humain.

L'architecture présentée permet d'aborder tant les aspects dialogiques que méta-dialogiques. On a vu que l'adaptation de la machine par rapport à son interlocuteur (profil, but, etc.) permet de rendre le DHM plus crédible et plus performant.

Un point intéressant du système est l'intégration de ressources déjà existantes (base de données de l'INRIA et outil morphologique de Ranx-Xérox notamment). Cela suit la tendance actuellement observée qui consiste à utiliser autant que possible, de façon modulaire, des ressources préexistantes, que l'on peut atteindre via le réseau Internet. Mais en contrepartie, 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.

Enfin, d'un point de vue ergonomique, une tentative d'amélioration de la présentation des résultats d'une requête a été menée. Nous avons intégré une vue en œil de poisson dans le système Halpin-Documentaire, et avons effectué des expériences avec des utilisateurs volontaires, afin d'évaluer ce mode de représentation d'une grande masse d'informations couplée à notre système de DHM. Les résultats montrent que, dans ce cas précis, le dialogue avec la machine était plus utilisé que la navigation à la souris sur la vue en forme de colline.

◆ Perspectives

Des améliorations devront être apportées quant à la robustesse du système lors d'erreurs et de mauvaises interprétations des énoncés de l'utilisateur. Une meilleure gestion des procédés anaphoriques, tout comme les ellipses ou les tropes, est à prendre en compte, car ils assurent la concision et la souplesse d'un dialogue en LN : « *Or, s'il est vrai que le traitement informatique de l'anaphore soulève des difficultés actuellement insolubles, il n'en reste pas moins vrai aussi que le recours à l'anaphore semble lié à l'effort fait dans la formulation, au souci de concision et de resserrement de la syntaxe.* » [MOREL 89].

Lorsqu'il subsistera des ambiguïtés, la machine sera toujours à même de dialoguer avec l'utilisateur, et de lui expliquer ce qu'elle n'arrive pas à saisir, plutôt que de faire un choix hasardeux : « *Les ambiguïtés par ellipses, anaphores, dans des messages écrits ou oraux, vont engendrer très souvent des demandes de reformulation de la part des interlocuteurs, avec ce que cela suppose de perte de temps. Soumis à un système automatique, un message ambigu est souvent rejeté, ou dans le meilleur des cas, le système fait des hypothèses dont il n'a nul moyen de vérifier la véracité.* » [PRINCE 96].

Pour aller plus loin, nous pourrions fournir la possibilité d'utiliser une reconnaissance vocale distante, et non plus seulement locale, comme c'est le cas avec IBM ViaVoice, que nous utilisons pour permettre un dialogue oral.

Le système Janus III, ou plus exactement, sa version française (Raphaël) [AKBAR 98], est actuellement en cours d'intégration dans le système Halpin, et sera disponible pour la prochaine version, de manière à s'orienter vers une complète prise en charge des ressources acoustiques par notre serveur.

Cela devrait permettre un accès au système par des postes clients ne possédant ni logiciel de synthèse, ni de reconnaissance vocale sur leur propre machine. Les problèmes d'attente, entraînant une rupture de l'aspect naturel du dialogue, que dénoncent certains utilisateurs, devraient donc pouvoir être surmontés d'ici quelque temps. De plus cela permettrait d'étendre la possibilité d'interagir de manière vocale, à tout type de machine (Mac, Unix), et non plus uniquement aux ordinateurs de type PC.

Notons que certains langages de développement orientés vers l'oral et l'utilisation de la langue naturelle commencent à voir le jour : JavaTM Speech API, par exemple, ou encore le langage VoxML de Motorola⁵², qui s'appuie sur le principe des balises du HTML, et devrait permettre à un usager du Web d'interagir avec un ordinateur, à la voix.

Du point de vue linguistique, l'intégration d'un thesaurus général français améliorerait considérablement la couverture de vocabulaire du système. Cela permettrait une plus grande compréhension des énoncés de l'utilisateur (le thesaurus indique par exemple que *journal* est une sorte de *document*). D'autre part, cela permettrait à la machine de varier ses répliques, en utilisant des synonymes, ou des locutions proches, de manière à assurer un dialogue vraisemblable, cohérent et le plus naturel possible.

Enfin, pour améliorer le dialogue, il serait peut-être intéressant de présenter dans les systèmes de dialogue un visage de synthèse animé, articulant de manière coordonnée la synthèse vocale prononcée par la machine.

⁵² « Motorola created the VoxML language in response to the need to provide an easier way to produce voice applications. The main goal of Motorola's VoxML effort is to offer a common approach and broadly supported platform for voice applications just as HTML provides for Web based applications » <http://www.voxml.com/voxml.html>.

L'agent August, proposée par [BELL et GUSTAFSON 99], et présenté sur la Figure 36 ci-dessous, dialogue en suédois avec ses interlocuteurs. Il est capable de reproduire des mimiques telles que l'interrogation, le doute, l'étonnement, etc.



Figure 36 : August, un agent animé, qui dialogue en suédois.

De même, Microsoft propose des agents animés que l'on intègre aux pages Web (cf. Figure 37), en insérant du code dans l'entête d'une page HTML. Ceci est censé améliorer l'interaction entre l'homme et la machine, puisque l'agent peut adopter des attitudes explicites (comme la surprise, l'explication, l'incompréhension, ...), se déplacer sur l'écran, désigner des parties de l'écran (à gauche, à droite, etc.). Ces MS-Agent [MICROSOFT] disposent de logiciels de synthèse et de reconnaissance vocale (en anglais pour le moment), qui s'installent sur la machine cliente, lors du chargement de la page Web.



Figure 37 : Quelques attitudes de Merlin (MS-Agent)

Toutes les recherches allant dans le sens de l'amélioration du DHM sont intéressantes, même s'il est vrai que la machine ne pourra pas remplacer l'homme dans toutes les situations : « (...) il s'agit de prendre les machines pour ce qu'elles sont, sans faire d'anthropomorphisme, car chacun préfère dialoguer avec des humains pour ce qui est de la conversation courante. Mais dans beaucoup de domaines, les machines jouent des rôles que les humains ne peuvent pas jouer : elles ne s'ennuient jamais quand elles répètent la même chose, elles ont une mémoire sûre et sans limite pratique, elle calculent plus sûrement, elles font moins d'erreurs. Or les limites à leurs usages, les appréhensions et difficultés des usagers, tiennent souvent aux difficultés de l'interaction. » [NICOLLE et LUZZATI 99].

Bibliographie

[AKBAR 98] AKBAR M., CAELEN J., *Parole et traduction automatique : le module de reconnaissance RAPHAEL*, COLLING-ACL'98, pp. 36-40, Montreal (Quebec), August 1998.

[ALLEN et LITMAN 87] ALLEN, J.F., LITMAN, D.J., *A Plan Recognition Model for Subdialogues in Conversations*, Cognitive Science 11, 2, 1987.

[ALLEN et PERRAULT 80] ALLEN J.F., PERRAULT, C.R., *Analyzing Intention in Utterances*, Artificial intelligence, n°15, 1980.

[ANDRÈS 95] ANDRÈS, M., *Étude des stratégies de prise de décision d'une machine dans le contexte d'un dialogue enfant-machine*, Thèse de doctorat d'informatique, Université de Caen, 1995.

[ANDRIEU 96] ANDRIEU, O., *Méthodes et outils de recherche sur l'Internet*, Eyrolles, Paris, 1996.

[ANGLESSY 89] ANGLESSY, N., *Diana : un système de dialogues naturels entre l'homme et la machine. Application aux services SNCF*. Thèse de doctorat d'informatique, Université Pierre et Marie Curie, Paris VI, 1989.

[APPELT 81] APPELT, D., *Planning natural language utterances to satisfy multiple goals*, Thèse de Ph.D., Stanford University, 1981.

[ASOH et al. 99] ASOH, H., MATSUI, T., FRY, J., ASANO, F., HAYAMIZU, S., *A Spoken Dialog System for a Mobile Office Robot*, Eurospeech'99, Budapest, pp. 1139-1142, 1999.

[BATEMAN et al. 95] BATEMAN, J. A.; HAGEN, E., STEIN, A. *Dialogue modelling for speech generation in multimodal information systems*, in P. Dalsgaard, et al. (Ed.), *ESCA Workshop on Spoken Dialogue Systems - Theories and Applications* (pp. 225-228). Aalborg, Denmark: ESCA/Aalborg University, 1995.

[BEAUFILS 96] BEAUFILS, A., *Approximation de réseaux de navigation hypermédias par des élèves de collège. Résultats d'une expérimentation et perspectives*. Troisième colloque Hypermédias et Apprentissage, Châtenay-Malabry, Mai 1996.

[BECKER et al. 92] BECKER, R., KUPPER, D., STROBEL, M., ROSNER, D., *A cooperative, natural language environmental information system*, Education and Society, R. Aiken (editor). Information Processing 92, volume II. Elsevier Science Publishers B.V. (North_Holland), 1992.

[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.

[BELL et GUSTAFSON 99] BELL, L., GUSTAFSON, J., *Interaction with an Animated Agent in a Spoken Dialogue System*, Eurospeech'99, Budapest, pp. 1159-1162, 1999.

[BELLIK et TEIL 92] BELLIK, Y., TEIL, D., *Multimodal Dialogue Interface*, WWDU'92, Berlin, 1992.

[BENNACEF et al. 96] BENNACEF, S., DEVILLERS, L., ROSSET, S., LAMEL, L., *Dialog in the RailTel telephone-based system*, In International Conference on Speech and Language Processing, pp. 550-553, Philadelphia, 1996.

[BILANGE 92] BILANGE, E., *Dialogue personne-machine : modélisation et réalisation informatique*, Collection Langue, raisonnement, calcul. Paris, Hermès, 1992.

[BLOMBERG et al. 93] BLOMBERG, M. CARLSON, R., ELENIUS, K., GRANSTROM, B., GUSTAFSON, J., HUNNICUTT, S., LINDELL, R., NEOVIUS, L., *An Experimental Dialogue System : Waxholm*, Eurospeech, pp. 1867-1870, Berlin, 1993.

[BOBROW et al 77] BOBROW, D., KAPLAN, J., NORMAN, D., THOMPSON, H., WINOGRAD, T., *GUS : a frame-driven dialog system*, In : *Artificial Intelligence* n. 8, pp. 155-173, 1977.

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

[BOUFAIDIA 93] BOUFAIDIA, M., *Génération d'interfaces et conduite du dialogue : le système GEODE*, Thèse de doctorat d'informatique, Université de Paris 6, 1993.

[BOUHAI et al. 99] BOUHAI, N. et SALEH, I.. *HyMapWeb : Système de gestion et de navigation hypertextuelles par profil*. Cinquième conférence Internationale Hypertextes & Hypermédiat - Septembre 1999, Paris.

[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.

[BRUN 98] BRUN, C., *A Terminology Finite-State Preprocessing for Computational LFG*. 36th International meeting of the Association for Computational Linguistics & 17th International Conference on Computational Linguistics, Montreal, Quebec, Canada, August 1998.

[BULLWINKLE 77] BULLWINKLE, C., *The semantic component of PAL: the personal assistant language understanding program*, note interne 141, Laboratoire d'intelligence artificielle, MIT, Massachusetts, USA, 1977.

[CAELEN 92] CAELEN, J. *Compte-rendu du « Workshop » Interfaces Homme-Machine multimodales*, GDR-PRC CHM sur le Dialogue, Dourdan, 1992.

[CARBERRY 90] CARBERRY, S., *Plan recognition in natural language*, MIT Press, Cambridge, Mass., 1990.

[CARBONELL et al. 83] CARBONELL, J.G., BOGGS, W.M., MAULDIN, M.L., ANICK, P. G., *The X-CALIBUR project: a natural language interface to expert systems*, 8th IJCAI, Karlsruhe, Germany, 1983.

[CARRÉ et al. 91] CARRÉ, R., DÉGREMONT, J.F., GROSS, M., PIERREL, J.M., SABAH, G., *Langage humain et machine*, Presses du CNRS, 1991.

[CATACH 85] CATACH, N., *Les listes orthographiques de base du français (LOB), Les mots les plus fréquents et leurs formes fléchies les plus fréquentes*, édition Nathan Recherche, 1984.

[CATINIS 98] CATINIS, L., *Etude de l'usage de la parole dans les interfaces multimodales*, Thèse de doctorat d'informatique, Université de Grenoble, 1998.

[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.

[CHANET 96] CHANET, C., *La demande dans le dialogue finalisé: de la surface linguistique aux représentations de l'action*. Thèse en sciences de l'information et de la communication, Grenoble, 1996.

[COLINEAU 97] COLINEAU, N., *Etude des marqueurs discursifs dans le dialogue finalisé*, Thèse de doctorat de sciences cognitives, Université Grenoble 1, 1997.

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

[COULON et KAYSER 86] COULON, D., KAYSER, D., *Informatique et langage naturel: Présentation générale des méthodes d'interprétation des textes écrits*, Revue TSI: Technique et Science Informatique, Vol. 5. n°2, 1986.

[COURSIL 92] COURSIL, J., Grammaire analytique du français contemporain, Essai d'intelligence artificielle et de linguistique générale, Thèse de doctorat, Université de Caen, 1992.

[CROFT et al. 90] CROFT W.B., BELKIN, N., BRUANDET, M-F, KUHLEN, R., OREN, T., *Hypertext and information retrieval : what are the fundamental concepts ?*, First European Conference on Hypertext , Versailles, 1990.

[DE MORI 94] DE MORI, R., *Apprentissage automatique pour l'interprétation sémantique*, XXèmes Journées d'étude sur la parole, Trégastel France, 1994.

[DEN OS et al. 99] DEN OS, E., BOVES, L., LAMEL, L., BAGGIA, P., *Overview of the ARISE Project*, Eurospeech'99, Budapest, pp. 1527-1530, 1999.

[DUTOIT et PAGEL 96] DUTOIT, T., PAGEL, V., *Le projet MBROLA : vers un ensemble de synthétiseurs vocaux disponibles gratuitement pour utilisation non-commerciale.*, XXI^{es} Journées d'Étude sur la Parole, Avignon, 1996.

[DUTOIT et al. 96] DUTOIT, T., PAGEL, V., PIERRET, N., BATAILLE, F., VAN DER VRECKEN, O., *The MBROLA Project: Towards a Set of High-Quality Speech Synthesizers Free of Use for Non-Commercial Purposes*, ICSLP'96, Philadelphia, vol. 3, pp. 1393-1396.

[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.

[ELAN] <http://www.elan.fr>

[FALZON 89] FALZON, P., *Ergonomie cognitive du dialogue*, Presses Universitaires de Grenoble, 1989.

[FISCHER et al. 94] FISCHER, M., MAIER, E., STEIN, A., *Generating Cooperative System Responses in Information Retrieval Dialogues*, In Proceedings of the International Workshop on Natural Language Generation (INLWG '94), Kennbunkport, Maine, pp. 207-216, 1994.

[FRECHET 88] FRECHET, A-L, *Analyse linguistique d'un corpus de dialogue Homme-Machine (oral finalisé)*, thèse de doctorat, Université de la Sorbonne nouvelle, Paris III, 1988.

[GAUSSIÉ et al. 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.

[GAUVAIN et al. 90] GAUVAIN, J.L., LAMEL, L.F., ESKENAZI, M., *Design Considerations and Text Selection for BREF, a large French read-speech corpus*, ICSLP-90, Kobe, Japan, 1990.

- [GERARD et NICOLLE 98] GERARD F., NICOLLE, A., *Bistro, un modèle de dialogue intégrant la manipulation de concepts*, Rencontre des Étudiants Chercheurs en Informatique pour le Traitement Automatique des Langues, RECTAL'98, Le Mans, septembre 1998.
- [GLASS et al. 95] GLASS, J., FLAMMIA, G., GOODINE, D., PHILLIPS, M., POLIFRONI, J., SAKAI, S., SENEFF, S., ZUE, V., *Multilingual Spoken-Language Understanding in the MIT Voyager System*, Speech Communication, 17, pp. 1-18, 1995.
- [GOLDSTEIN et ROBERTS 79] GOLDSTEIN, I., ROBERTS, B., *Using frame in scheduling*, in Artificial intelligence, a MIT perspective, Volume 1, Winston et Brown, MIT Press, Cambridge, Mass., USA, 1979.
- [GOUGENHEIM et al. 56] GOUGENHEIM, G., MICHEA, G.R., RIVENC, P., SAUVAGEOT, A., *L'élaboration du français élémentaire*, Didier, Paris, 1956.
- [GRAU et al. 94] GRAU, B., SABAH, G., VILNAT, A., *Pragmatique et dialogue homme/machine* In : *Technique et Science Informatique*, vol. 13, n. 1, 1994.
- [GRIMALDI 93] GRIMALDI, L., *Dialogue pédagogique en langue naturelle dans un système intelligent de communication de connaissances : application à la résolution d'exercices au jeu d'échecs*. Thèse de doctorat d'informatique, Université Paris 6, 1993.
- [GUEDON 96] GUEDON, J-C, *La planète cyber*, Canal + Gallimard, Mars 1996.
- [GUYOMARD et al. 84] GUYOMARD, M., SIROUX, J., TRILLING, L., *Le dialogueur : un intermédiaire entre l'utilisateur et l'application*, Actes du séminaire GRECO Dialogue Homme-Machine, Nancy, 1984.
- [GUYOMARD et al. 95] GUYOMARD, M., NERZIC, P., SIROUX, J., *Plans, métaplans et dialogue*, Actes des journées de la communication parlée et dialogue, ENST, 1995.
- [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.
- [HUANG et al. 99] HUANG, C., XU, P., ZHANG, X., ZHAO, S., HUANG, T., XU, B., *LODESTAR: A Mandarin Spoken Dialogue System for Travel Information Retrieval*, Eurospeech'99, Budapest, pp. 1159-1162, 1999.
- [JACQUET 97] JACQUET, D., *L'utilisateur en dialogue Homme-Machine*, dans *Apprentissage par l'interaction*, K. Zreik, Europia Productions, Paris, 1997.
- [LAMEL et al. 91] LAMEL, L.F., GAUVAIN, J.L., ESKENAZI, M., *BREF, a Large Vocabulary Spoken Corpus for French* Proc. EUROSPEECH'91, Genoa, Italy, 1991.

[LAU et al. 97] LAU, R., FLAMMIA, G., PAO, C., ZUE, V., *WebGalaxy –Intergrating Spoken Language and Hypertext Navigation*, Eurospeech, Rhodes Greece, 1997.

[LE MOIGNE 86] LE MOIGNE, J.L., *Genèse de quelques nouvelles sciences : de l'intelligence artificielle aux sciences de la cognition, Intelligences des mécanismes - Mécanismes de l'intelligence*, Fayard, 1986.

[LE NY 79] LE NY, J.F., *La sémantique psychologique*, Presse universitaire, Paris, 1979.

[LEHUEN 97] LEHUEN J., *Un modèle de dialogue dynamique et générique intégrant l'acquisition de sa compétence. Le système COALA*. Thèse de doctorat, Université de Caen, Juin 1997.

[LEMEUNIER 99] LEMEUNIER, T., *La modélisation des attentes dans le système AMI*, Actes du colloque RéciTAL'99, Institut d'Etudes Scientifiques de Cargese (Corse), pp. 439-443, 1999.

[LIFE et al. 96] LIFE, A., SALTER, I., TEMEM, J.N., BERNARD, F., ROSSET, S., BENNACEF, S., LAMEL, L., *Data collection for the Mask kiosk: WOz vs prototype system* In International Conference on Speech and Language Processing, pp. 1672-1675, Philadelphia, 1996.

[LITMAN 85] LITMAN, D.J., *Plan recognition and discourse analysis : an integrated approach for understanding dialogues*, Thèse de l'université de Rochester, 1985.

[LOPEZ-COZAR et al. 99] LOPEZ-COZAR, R., RUBIO, A. J., GARCIA, P., SEGURA, J. C., *A New Word-Confidence Threshold Technique to Enhance the Performance of Spoken Dialogue Systems*, Eurospeech'99, Budapest, pp. 1395-1398, 1999.

[LUZZATI 89] LUZZATI, D., *Recherches sur le dialogue homme-machine : modèles linguistiques et traitements automatiques*, Thèse de doctorat, Université de la Sorbonne, 1989.

[LUZZATI 95] LUZZATI, D., *Le dialogue verbal homme-machine, étude de cas*. Collection Sciences cognitives. Masson, 1995.

[MANN et MOORE 81] MANN, W., MOORE, J., *Computer generation of multiparagraph English text*, *American journal of computational linguistics*, 7, 1, pp. 17-29, 1981.

[MATROUF et al. 87] MATROUF, A., NEEL, F., MARIANI, J., *Système de dialogue orienté par la tâche : une application en avionique*, 16^{ème} Journée d'Etude sur la Parole, Société Française d'Acoustique, Hammamet, 1987

[MAYFIELD et al. 95] MAYFIELD, L., GAVALDA, M., SEO, Y-H., SUHM, B., WARD, W., WAIBEL, A., *Parsing real input in Janus: a concept based approach to spoken language translation*, Proceedings of TMI 1995.

[MC COY 85] MC COY, K.F., *The role of perspective in responding to property misconceptions*, Actes 9^o IJCAI, Los Angeles, pp. 791-793, 1985.

[MC KEOWN 82] MC KEOWN, K., *Generating natural language in text in response to questions about database structure*, Thèse de PhD., Université de Pennsylvanie, Philadelphie, 1982.

[MEIYE et al. 98] MEIYE, J.P., ROUILLARD, J., VAUFREDAZ, D., *Webcompletion - protocole de normes associatives sur Internet*. Ecole thématique CNRS - Sémantique, Caen (Asnelles-sur-mer), 1998.

[MELONI et LANGLAIS 96] MELONI, H., LANGLAIS, P., *Prosodie et Reconnaissance de la Parole*, in « Fondements et perspectives en traitement automatique de la parole », coordinateur Henri Méloni, AUPELF-UREF, 1996.

[MICROSOFT] <http://msdn.microsoft.com/msagent>

[MOREL 85] MOREL, M-A., *Analyse linguistique d'un corpus d'oral finalisé*, Rapport du Gréco « Communication parlée », 1985.

[MOREL 88] MOREL, M-A., *Analyse linguistique d'un corpus de dialogues homme-machine*, Tome 1, premier corpus : centre de renseignement SNCF à Paris, Publications de la Sorbonne Nouvelle, 1988.

[MOREL 89] MOREL, M-A., *Analyse linguistique d'un corpus de dialogues homme-machine*, Tome 2, deuxième corpus : centre d'information et d'orientation de l'université de Paris V, Publications de la Sorbonne Nouvelle, 1989.

[MORIN 84] MORIN, E., *L'extraordinaire problème : le cerveau et l'esprit*, AFCET-Interfaces, n°25, 1984 ; texte repris dans *La connaissance de la connaissance – la Méthode*, Tome III, Paris, édition du Seuil, 1986.

[MORIN 94] MORIN, P., *Un Système de Dialogue Homme-Machine Multimodal pour Applications Finalisées à Forte Composante Orale*, Thèse d'informatique, Université Henri Poincaré, Nancy I, Vandoeuvre-lès-Nancy, 1994.

[NERZIC 93] NERZIC, P., *Erreurs et échecs dans le dialogue oral homme-machine : détection et réparation*. Thèse de doctorat, Université de Rennes, 1993.

[NICOLLE et LUZZATI 99] NICOLLE, A., LUZZATI, D., *La langue dans l'interaction personnes/machines*, Atelier thématique de TALN'99, Cargèse (Corse), 1999.

[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 94] NIGAY, L., *Conception et modélisation logicielles des systèmes interactifs : application aux interfaces multimodales*. Thèse d'informatique de l'université de Grenoble I, Janvier 1994.

[NOUALI 97] NOUALI, O., SIDI BOUDEDINE, S.R., *Système de génération automatique d'interfaces pour l'interrogation des bases de données en langue naturelle*. Premières Journées Scientifiques et Techniques du Réseau Francophone de l'Ingénierie de la Langue de l'AUPELF-UREF, Avignon, Avril 1997.

[ORION 99] <http://www.gate.cnrs.fr/~zeiliger/Orion99.doc>

[PATERAS et al. 99] PATERAS, C., CHAPADOS, N., KWAN, R., LAVOIE, D., TREMBLAY, R., *A Mixed-Initiative Natural Dialogue System for Conference Room Reservation*, Eurospeech'99, Budapest, pp. 1275-1278, 1999.

[PEARL 84] PEARL, J., *Heuristics: Intelligent Search Strategies for Computer Problem Solving*, vol. I. Addison-Wesley, 1984.

[PECKHAM 91] PECKHAM, J., *Speech Understanding and Dialog over Telephone : an overview of the ESPRIT SUNDIAL Project*, 2nd European Conference on Speech Communication and Technology, Genève, pp. 1469-1472, 1991.

[PIERACCINI et al. 91] PIERACCINI, R., *Stochastic Representation of Conceptual Structure in the ATIS Task*, in DARPA 1991.

[PIERACCINI et al. 97] PIERACCINI, R., LEVIN, E., ECKERT, W., *AMICA : the AT&T Mixed Initiative Conversational Architecture*, Eurospeech, Rhodes Greece, 1997.

[PIERREL 87] PIERREL, J.M., *Dialogue oral homme-machine*, Hermès, Paris, 1987.

[PIERREL et ROMARY 97] PIERREL, J.M., ROMARY, *Quelles références dans les dialogues homme-machine ?*, In SABAH, G., VIVIER, J., VILNAT, A., PIERREL, J.M., ROMARY, L., NICOLLE, A., *Machine, langage et dialogue*, Figures de l'interaction, L'Harmattan, 1997.

[PITKOW et KEHOE 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.

[POLITY et al. 90] POLITY, Y., FRANCONY, J-M., PALERMITI, R., FALZON, P., KAZMA, S., *Recueil de dialogues homme-machine en langue naturelle écrite*, Les Cahiers du CRISS, n° 17, 1990.

[PONTON 96] PONTON, C., *Génération de textes en langue naturelle. Essai de définition d'un système noyau*, thèse de doctorat, Université Stendhal, Grenoble III, 1996.

- [POSTMAN 79] POSTMAN, N., *Enseigner c'est résister*, Paris, 1981 (1979), Le Centurion, Trad. J. Chambert et D-J., Piveteau., in ROMAN, J., *Chronique des idées contemporaines*. Bréal, 1995.
- [PRINCE 96] PRINCE, V., *Vers une informatique cognitive dans les organisations*, Collection Sciences cognitives. Masson, 1996
- [RAMAUX 95] RAMAUX, N., *Acquisition Automatique et Polysémie en Langage Naturel*, rapport de DEA, laboratoire HEUDIASYC, groupe LCS, Compiègne, 1995.
- [RASTIER 87] RASTIER, F., *Dialogue Homme-machine et représentation de l'interlocuteur*, Technologos, volume Printemps 87, numéro 4, p. 35-53.
- [RENAULT 95] RENAULT, *PROMETHEUS : fluidité et sécurité routière*, Bulletin Technique, Véhicules Renault, 74, 2-33, 1995.
- [ROUILLARD 99] ROUILLARD, J., *Le système HALPIN : Hyperdialogue avec un Agent en Langage Proche de l'Interaction Naturelle*, TALN'99, Cargèse, Corse, 1999.
- [ROUILLARD 98] ROUILLARD, J., *Hyperdialogue Homme-Machine sur le World Wide Web : Le système HALPIN*, Colloque international Ergonomie et Informatique avancée ERGOIA'98, Biarritz, 1998.
- [ROUILLARD et CAELEN 98] ROUILLARD, J., 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, 1998.
- [ROUSSANALY 88] ROUSSANALY , A. *DIAL : la composante dialogue d'un système de communication orale homme-machine finalisée en langage naturel*, Thèse de doctorat, Université Nancy 1, 1988.
- [ROUSSANALY et PIERREL 92] ROUSSANALY , A., PIERREL, J.M., *Dialogue oral homme-machine en langage naturel : le projet DIAL*, Rapport du CRIN n° 92-R-071, 1992.
- [SABAH 89] SABAH, G., *L'intelligence artificielle et le langage. Processus de compréhension*, Vol. 2, Hermès, Paris, 1989.
- [SABAH et al. 97] SABAH, G., VIVIER, J., VILNAT, A., PIERREL, J.M., ROMARY, L., NICOLLE, A., *Machine, langage et dialogue*, Figures de l'interaction, L'Harmattan, 1997.
- [SABAH 97] SABAH, G., *Apprentissage et traitement automatique des langues*, dans *Apprentissage par l'interaction*, K. Zreik, Europa Productions, Paris, 1997.

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

[SAMIER et SANDOVAL 98] SAMIER, H., SANDOVAL, V., *La recherche intelligente sur l'Internet, outils et méthodes*, Hermès, 1998.

[SASAJIMA et al. 99] SASAJIMA, M., YANO, T., KONO, Y., *EUROPA: A Generic Framework for Developing Spoken Dialogue Systems*, Eurospeech'99, Budapest, pp. 1163-1166, 1999.

[SIROUX et al. 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.

[SIROUX et al. 95] SIROUX, J., GUYOMARD, M., JOLLY, Y., MULTON, F., REMONDEAU, C., *Speech and Tactile-Based Géoral System*, In *proceedings of 4th European Conference on Speech Communication and Technology*, pp. 1943-1946, Madrid, Espagne, 1995.

[SMAÏL 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.

[SOUVAY 92] SOUVAY, G., *DIAPASON : un environnement de développement pour l'intégration d'une entrée vocale dans les applications de type commande de machine*, Thèse d'informatique, Université Henri Poincaré, Nancy I, Vandoeuvre-lès-Nancy, 1992.

[STEIN et MAIER 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.

[STEIN et al. 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.

[TALEB 96] TALEB, L., *Recherches sur l'Interaction Homme-Machine : les écarts dans le dialogue informatif finalisé*, Thèse de doctorat, Université Paris 3, 1996.

[TRIGANO 94] TRIGANO, P., *Des Bases de Données et de Connaissances aux Interfaces en Langue Naturelle: un Enrichissement Réciproque*, Habilitation à diriger des recherches, Université de Paris XIII (Paris-Nord), Villetaneuse, 1994.

[VEILEX 85] VEILEX, F. *Approche expérimentale des processus humains de compréhension en vue d'une indexation automatique des résumés scientifiques. Application à un corpus de géologie*. Thèse de doctorat, Université Grenoble II, Cahier du CRISS, n°6, 1985.

[VERNANT 94] VERNANT, D., *Dialectique, forme dialogale et dialogique* In : Actes du colloque Le Dialogique, sur les formes philosophiques, littéraires, linguistiques et cognitives du dialogue, Le Mans, 1994.

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

[VILLASEÑOR-PINEDA 97] VILLASEÑOR-PINEDA, L., *Contribution à l'apprentissage dans le dialogue homme-machine*, Thèse de doctorat d'informatique, Université de Grenoble, 1999.

[VILNAT 97] VILNAT, A., *Quels processus pour les dialogues homme-machine ?* In SABAH, G., VIVIER, J., VILNAT, A., PIERREL, J.M., ROMARY, L., NICOLLE, A., *Machine, langage et dialogue*, Figures de l'interaction, L'Harmattan, Paris, 1997.

[VIVIER et NICOLLE 97] VIVIER, J., NICOLLE, A., *Questions de méthode en dialogue homme-machine : l'expérience Compèrobot*, In SABAH, G., VIVIER, J., VILNAT, A., PIERREL, J.M., ROMARY, L., NICOLLE, A., *Machine, langage et dialogue*, Figures de l'interaction, L'Harmattan, 1997.

[WATZLAWICK 80] WATZLAWICK, P., *Le changement. Eléments de communication thérapeutique*. Paris : Seuil. (Réédition en collection "Points" , n°186, 1986).

[WEIZENBAUM 66] WEIZENBAUM, J., *Éliza - a computer program for the study of the natural language communication between man and machine*. . CACM n°9, 1966.

[WRETO et CAELEN 89] WRETO, J., CAELEN, J., *ICP-Draw. Rapport final du projet ESPRIT Multiworks n°2105, 1989*.

[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.

[ZUE et al. 94] ZUE, V., SENEFF, S., POLIFRONI, J., PHILLIPS, M., PAO, C., GODDEAU, D., GLASS, J., BRILL, E., *PEGASUS: A Spoken Language Interface for On-Line Air Travel Planning*, Speech Communication, 15, pp. 331-340, 1994.

[ZUE 97] ZUE, V., *Conversational Interfaces: Advances and Challenges*, KN speech, Eurospeech, Rhodes Greece, 1997.

Table des illustrations

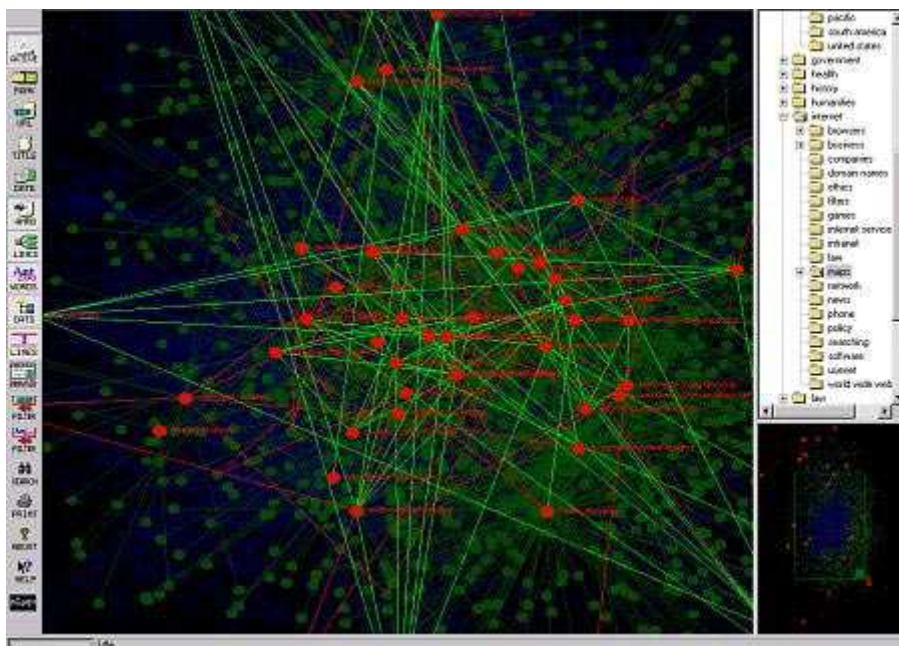
<i>Figure 1 : Le premier "browser" connu (Paris, 1588)</i>	11
<i>Figure 2 : Méthodologie mise en œuvre pour la conception de notre système de DHM</i>	14
<i>Figure 3 : Code SQL équivalent à la requête en LN « Which sewage treatment plants have a chemical purifying stage ? »</i>	17
<i>Figure 4 : Quelques systèmes de DHM de type U1</i>	34
<i>Figure 5 : Quelques systèmes de DHM de type U2</i>	35
<i>Figure 6 : Quelques systèmes de DHM de type U3</i>	36
<i>Figure 7 : Quelques systèmes de DHM positionnés sur les plans U1, U2 ou U3</i>	37
<i>Figure 8 : Avis des usagers sur le catalogue de recherche des documents de la médiathèque</i>	46
<i>Figure 9 : Recherches selon un thème ou un domaine</i>	47
<i>Figure 10 : Modes de recherche de documents des usagers (question à choix multiple)</i>	48
<i>Figure 11 : Pertinence des informations données par la machine, selon les utilisateurs</i>	49
<i>Figure 12 : Interruption volontaire de recherche par les utilisateurs en cas d'incompréhension de la machine</i>	50
<i>Figure 13 : Souhait d'un dialogue plus interactif avec la machine</i>	50
<i>Figure 14 : Hypothèse d'utilisation d'un ordinateur avec reconnaissance de la parole pour la recherche d'ouvrages à la médiathèque ?</i>	51
<i>Figure 15 : Un système générique d'hyperdialogue pour le WWW</i>	58
<i>Figure 16 : Architecture du système Halpin-Recueil</i>	65
<i>Figure 17 : Carte de France des connexions au système Halpin-Recueil</i>	70
<i>Figure 18 : Tours de parole dans le corpus Halpin-Recueil Analyse du corpus Halpin-Recueil</i>	75
<i>Figure 19 : Fréquence d'apparition des 100 premières occurrences du corpus Halpin-Recueil</i>	77
<i>Figure 20 : Le concept du « thème »</i>	78
<i>Figure 21 : Principe de fonctionnement d'une application utilisant des technologies vocales</i>	80
<i>Figure 22 : Le concept « acceptation » (autre que oui)</i>	84
<i>Figure 23 : Vocabulaire employé dans les dialogues Homme-Machine du corpus Halpin-Recueil</i>	87
<i>Figure 24 : Fréquences relatives de pronoms en fonction des différentes vagues d'expériences</i>	93
<i>Figure 25 : L'utilisateur dialogue avec le système Halpin-Documentaire.</i>	100
<i>Figure 26 : Arbre de concepts utilisé pour la reconnaissance de concepts dans les énoncés</i>	104
<i>Figure 27 : Gestion des événements du système Halpin-Documentaire</i>	117
<i>Figure 28 : Architecture du système Halpin-Documentaire</i>	125
<i>Figure 29 : Détection de parole en dispositif « main libre »</i>	133
<i>Figure 30 : L'interface du système Halpin-Documentaire pour un DHM multimodal sur le Web</i>	137
<i>Figure 31 : Interaction traditionnelle avec le formulaire de recherche de l'INRIA</i>	147
<i>Figure 32 : Utilité du système Halpin par rapport à un formulaire traditionnel du Web</i>	149

<i>Figure 33 : Temps relatif nécessaire pour fournir une interaction dialoguée sur le Web</i>	151
<i>Figure 34 : Représentation graphique de Livetopics pour le moteur de recherche Altavista</i>	151
<i>Figure 35 : Interface du système Halpin-Documentaire, intégrant une vue en œil de poisson, dite « en colline »</i>	152
<i>Figure 36 : August, un agent animé, qui dialogue en suédois.</i>	168
<i>Figure 37 : Quelques attitudes de Merlin (MS-Agent)</i>	168
<i>Figure 38 : Une vue générale de l'écran de NESTOR</i>	210
<i>Figure 39 : Interface du système MERIT, utilisant le modèle de dialogue COR</i>	213
<i>Figure 40 : Le modèle COR (Conversational Roles)</i>	214
<i>Figure 41 : Sous-réseau détaillé de l'offre (B,A) dans le modèle COR</i>	215
<i>Tableau 1 : Exemple de lemmatisation de « de quoi parle le livre ? »</i>	65
<i>Tableau 2 : Exemple de lemmatisation de « quelle est l'année, alors ? »</i>	66
<i>Tableau 3 : Statistiques à propos du nombre de questions (homme) et de réponses (machine) par dialogue</i>	74
<i>Tableau 4 : Fréquence des 25 premières occurrences du corpus Halpin-Recueil</i>	78
<i>Tableau 5 : Quelques concepts associés aux critères de recherche</i>	83
<i>Tableau 6 : Quelques termes du vocabulaire employé par l'homme et par la machine</i>	85
<i>Tableau 7 : Fréquence d'apparition des verbes les plus utilisés par les utilisateurs</i>	86
<i>Tableau 8 : Erreur de lemmatisation : titre vs titrer</i>	86
<i>Tableau 9 : Fréquence d'apparition des noms les plus utilisés par les utilisateurs</i>	87
<i>Tableau 10 : Les concepts d'ouverture et de clôture de dialogue</i>	89
<i>Tableau 11 : Fréquence et rang des pronoms du corpus Halpin-Recueil</i>	92
<i>Tableau 12 : Quelques abréviations anglaises courantes pour les dialogues sur Internet</i>	95
<i>Tableau 13 : Les smileys les plus utilisés sur Internet</i>	96
<i>Tableau 14 : Exemples d'adaptation de la réponse en fonction des critères de recherche</i>	113
<i>Tableau 15 : Fiche de traitement de l'événement « Affiner »</i>	118
<i>Tableau 16 : Adaptation du dialogue selon le contexte et les connaissances de la machine</i>	122
<i>Tableau 17 : Analyse statistique du temps nécessaire pour une interaction Homme-Machine</i>	150
<i>Tableau 18 : Opinion des usagers à propos du système utilisé</i>	156
<i>Capture 1 : Interface du système Halpin-Recueil</i>	67
<i>Capture 2 : Interface du système Halpin-Recueil sur le Web : l'utilisateur a posé une question dans la zone de dialogue et la machine a répondu dans la zone de résultats</i>	68
<i>Capture 3 : Hyperdialogue entre l'homme et la machine</i>	69
<i>Capture 4 : Un exemple de dialogue avec le système Halpin-Recueil</i>	72
<i>Capture 5 : Exemple de dialogue avec plaisanterie</i>	73
<i>Capture 6 : Un exemple d'échec d'aide à la résolution de la tâche</i>	73
<i>Capture 7 : Un exemple d'échec dialogique avec le système Halpin-Recueil</i>	74
<i>Capture 8 : L'utilisateur n'exprime pas que des demandes</i>	81
<i>Capture 9 : L'utilisateur s'exprime avec la machine comme avec un interlocuteur humain</i>	81
<i>Capture 10 : L'utilisateur utilise des informations données par la machine</i>	82
<i>Capture 11 : L'utilisateur anticipe des informations données par la machine</i>	83

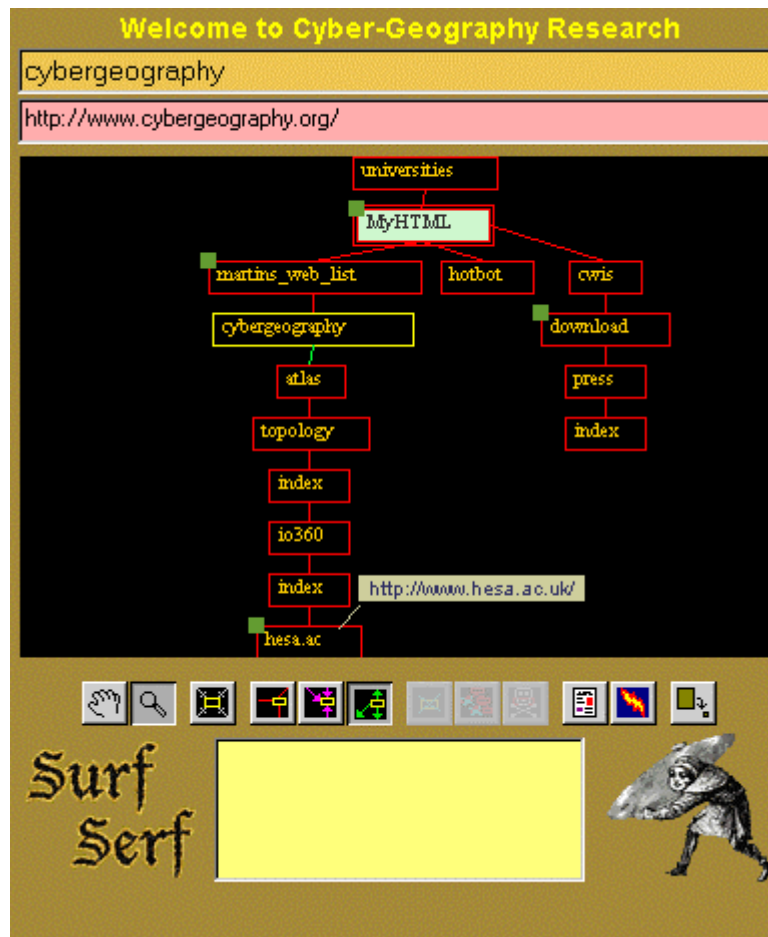
<i>Capture 12 : Exemples d'ouverture de dialogue dans le corpus Halpin-Recueil</i>	<i>90</i>
<i>Capture 13 : L'utilisateur doute des informations données par la machine</i>	<i>90</i>
<i>Capture 14 : Quelques énoncés des usagers comportant le mot "il", (corpus Halpin-Recueil).....</i>	<i>93</i>
<i>Capture 15 : Utilisation de mots contractés et de smileys dans un DHM sur Internet</i>	<i>95</i>
<i>Capture 16 : Exemple de métadialogue : l'ordinateur explique pourquoi il interrompt le dialogue relatif à la tâche.</i>	<i>112</i>
<i>Capture 17 : Reconnaissance de la parole avec envoi indirect</i>	<i>134</i>
<i>Capture 18 : Reconnaissance de la parole avec envoi direct</i>	<i>134</i>
<i>Capture 19 : Exemple N°1 d'un dialogue tiré du corpus Halpin-Documentaire.....</i>	<i>143</i>
<i>Capture 20 : Exemple N°2 d'un dialogue tiré du corpus Halpin-Documentaire.....</i>	<i>146</i>
<i>Capture 21 : Exemple N°3 d'un dialogue tiré du corpus Halpin-Documentaire.....</i>	<i>146</i>
<i>Capture 22 : Exemple N°4 d'un dialogue tiré du corpus Halpin-Documentaire.....</i>	<i>148</i>
<i>Capture 23 : Exemple de manque de robustesse dans le dialogue.....</i>	<i>160</i>
<i>Capture 24 : Interface du navigateur NESTOR lors d'une utilisation du système HALPIN.....</i>	<i>212</i>

Annexes

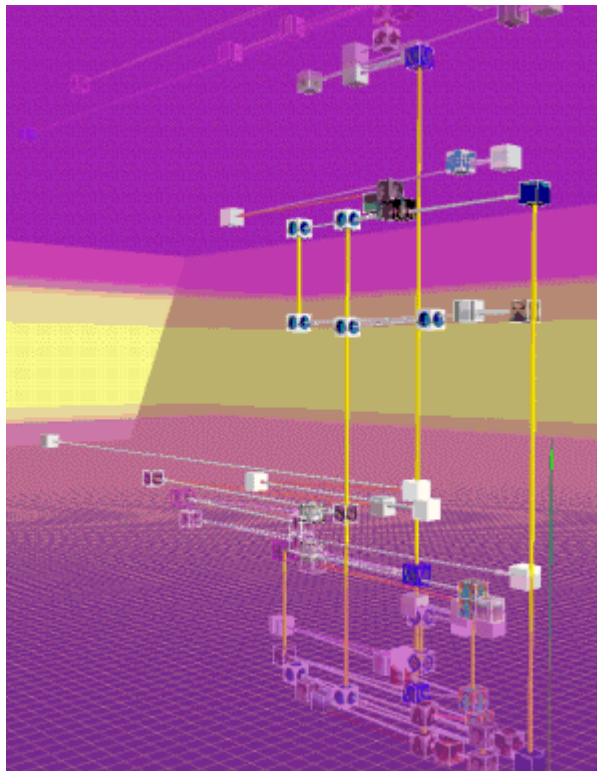
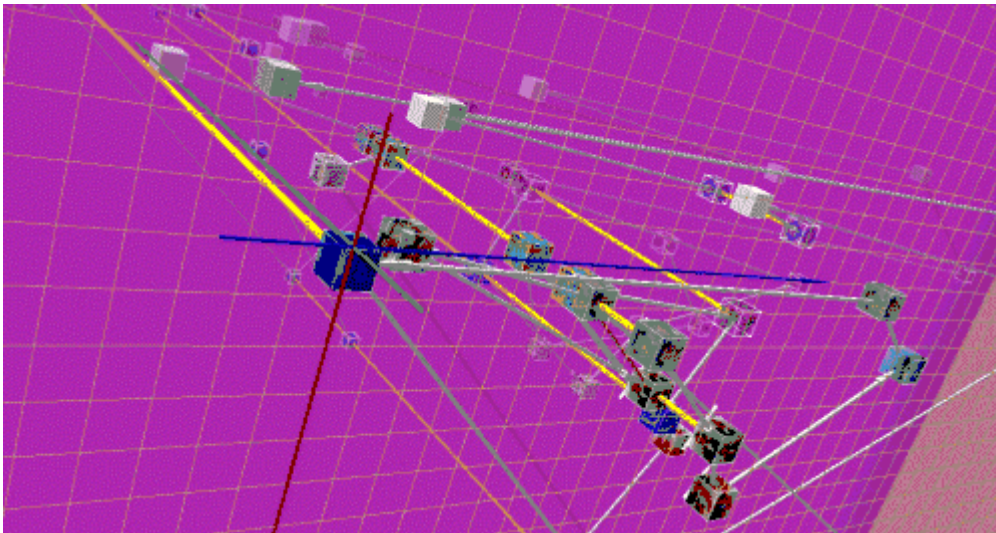
Annexe 1 : Quelques outils de navigation cartographique pour le Web



Internet Cartographer d'Inventix Software (<http://www.inventix.com>) est une application qui fonctionne en parallèle d'un navigateur classique, et qui cartographie toutes les pages visitées par l'utilisateur. Cette carte peut ensuite être interrogée pour la recherche de sites Web.



SurfSerf est une autre application d'aide à la navigation qui produit un graphe en deux dimensions du parcours de l'utilisateur sur le Web. Il présente également la structure des sites visités et la manière dont les pages sont liées.



Ces deux images montrent le logiciel WebPath ("*... a tool that unobtrusively visualises a user's trail as they browse the Web*"), développé par Emmanuel Frécon, chercheur dans le groupe d'environnements collaboratifs distribués au SICS.

http://www.sics.se/~emmanuel/home_prof.html.

Pour plus de détails, voir : FRECON E. et SMITH G., (1998), *WebPath - A three-dimensional Web History*, IEEE Symposium on Information Visualization (InfoVis '98), Chapel Hill. NC, USA, 1998.

Annexe 2 : Le questionnaire de l'enquête d'usage à la médiathèque IMAG

ENQUETE MEDIATHEQUE

Etat civil :

- * Age : ans
- * Sexe : M F
- * Nationalité : Française Autre * Langue maternelle :
- * Français écrit : Courant Moyen Faible
- * Français oral : Courant Moyen Faible

Etudes en cours :

- * Diplôme préparé :
- 2^{ème} Cycle : 3^{ème} Cycle :
- Chercheur : Autre :

A propos de la médiathèque :

• Fréquentation de la médiathèque :

- Tous les jours
- 1 à 2 fois /semaine
- 1 à 2 fois /quinzaine
- 1 à 2 fois / mois

• Estimez-vous avoir été bien informé des moyens disponibles pour la recherche de documents à la médiathèque :

- Très bien
- Assez bien
- Pas Assez

• Comment effectuez-vous votre recherche de documents :

- D'abord hors médiathèque (savoir ce que l'on veut en arrivant sur place)
- Par tâtonnement dans la médiathèque
- Par connaissance du catalogue (ex : je sais que la revue R parle de ...)
- Envoyé ici par une autre bibliothèque
- Bibliographie donnée par un professeur
- Bibliographie trouvée dans un autre ouvrage
- Bibliographie trouvée sur le Web
- Avec l'aide d'une bibliothécaire de la médiathèque

• Utilisation des terminaux de recherche :

- Toujours
 De temps en temps : Combien :
 Jamais ; pourquoi :

• Comment trouvez-vous le catalogue de recherche accessible par le web:

- Suffisant
 Insuffisant
et Pourquoi :.....
 Facile d'accès Accès moyen Difficile d'accès

• Quels sont selon vous les points à améliorer à ce catalogue :

- Information pour savoir si un ouvrage est disponible ou pas
 Réservation
 Sélection d'un champs (auteur, mot-clé du titre etc.)
 Autre :

• Vous trouvez la disposition géographique des ouvrages :

- Agréable Pas agréable
 Fonctionnelle Pas fonctionnelle
 Facilement « mémorisable » Pas facilement « mémorisable »

• Le code couleur utilisé pour repérer l'emplacement des ouvrages vous semble :

- Approprié
 Peu approprié
et Pourquoi :.....

L'utilisez-vous :

- Toujours
 De temps en temps
 Jamais

• Pouvez-vous donner quelques unes de vos attentes concernant la médiathèque ?

ex : Bouquin sur la linguistique

ex : Meilleure information par rapport aux ouvrages non trouvés dans les rayons

Meilleure Interface :

- L'interface proposée pour l'accès au catalogue vous gêne-t-elle actuellement lorsque vous effectuez une recherche :

- Non, pas du tout
 Un peu
 Oui, beaucoup

Et pourquoi :

- Vous arrive-t-il d'effectuer des recherches sans vraiment pouvoir définir ce que vous recherchez (ex : je veux connaître le nom d'un langage informatique utilisable sur le WWW, mais je ne connais pas *a priori*, Ni Perl, Ni Java, Ni le CGI) :

Non, jamais

Ca peut m'arriver

Oui, souvent

- Vous arrive-t-il d'interrompre vos recherches lorsque vous vous rendez compte que l'ordinateur ne comprend pas ce que vous voulez :

Oui

Non

- Vous arrive-t-il de souhaiter que le dialogue avec la machine soit plus interactif (ex : pouvoir lui dire «non, ce n'est pas cela que je cherche» ou «je veux plus de détails») :

Oui

Non

- D'une manière générale estimez-vous que les informations données par la machine sont :

Plutôt riches

Plutôt pauvres

- Globalement, estimez-vous que vous retrouvez l'information que vous recherchez :

Rapidement

Lentement

- Utiliserez-vous un ordinateur avec reconnaissance de la parole pour la recherche de vos ouvrages à la médiathèque :

Volontiers

Peut-être

Sûrement pas

Et pourquoi :

URL pour voir ces résultats sur le Web : [Http://hermes.imag.fr/~rouillar/mediatheque/enquete.html](http://hermes.imag.fr/~rouillar/mediatheque/enquete.html)

Annexe 3 : Utilisation d'interfaces de DHM oral dans le domaine militaire

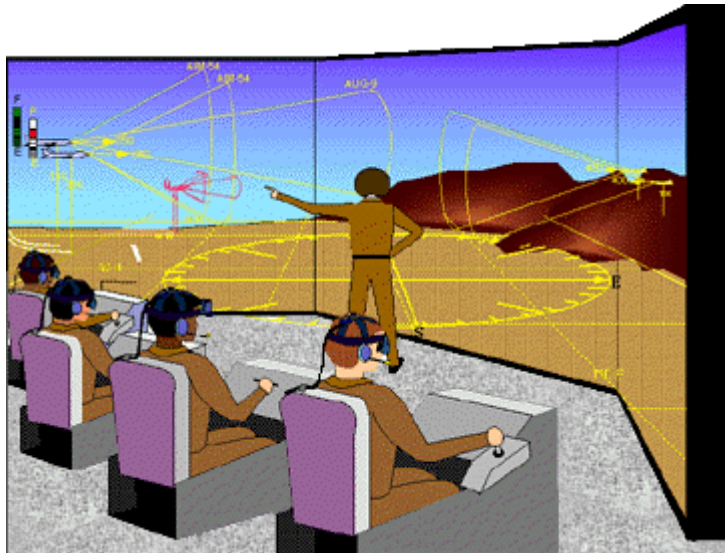
Source : <http://www.rti.org/vr/w/Nlp.cfm>



Sgt. Pike conversing with the Advanced Maintenance Assistant and Trainer (AMAT) while troubleshooting the laser range finder in the M1A1 Abrams tank.

COMMANDTALK : SPOKEN LANGUAGE INTERFACE TO THE LEATHERNET SYSTEM

Source : <http://www.ai.sri.com/~lesaf/brochure.html>



GOAL

Battlefield simulation is an area of growing importance for all military services. Giving commanders a natural way to control simulated forces remains a challenge, however. SRI International is addressing this problem by developing CommandTalk, a spoken-language interface that will allow commanders to use natural, spoken English commands to control simulated forces. This technology will reduce costs by providing more efficient training and will improve readiness by providing more realistic training.

Annexe 4 : Le test de Turing

Alan Turing présente son jeu de l'imitation (Test de Turing) de la manière suivante :

Je propose de considérer la question : les machines peuvent-elles penser ?

Il faudrait commencer par définir le sens des termes *machine* et *penser*. Les définitions pourraient être conçues de manière à refléter autant que possible l'utilisation normale des mots, mais cette attitude est dangereuse. Si on doit trouver la signification des mots *machine* et *penser* en examinant comment ils sont communément utilisés, il est difficile d'échapper à la conclusion que la signification de la question : les machines pensent-elles ? et la réponse à cette question, doivent être recherchées dans une étude statistique telle que le sondage d'opinion. Mais cela est absurde. Au lieu de m'essayer à une telle définition, je remplacerai la question par une autre, qui lui est étroitement liée et qui est exprimée en des termes relativement non ambigus.

Le problème reformulé peut être décrit dans les termes d'un jeu que nous appellerons: le jeu de l'imitation. Il se joue à trois, un homme (A), une femme (B), et un interrogateur (C) qui peut être de l'un ou l'autre sexe. L'interrogateur se trouve dans une pièce à part, séparé des deux autres. L'objet du jeu, pour l'interrogateur, est de déterminer lequel des deux est l'homme et lequel est la femme. Il les connaît sous les appellations X et Y et, à la fin du jeu, il doit déduire soit que : X est A et Y est B, soit que X est B et Y est A. L'interrogateur peut poser des questions à A et à B de la manière suivante :

C : « X peut-il ou peut-elle me dire, s'il vous plaît, quelle est la longueur de ses cheveux ? ».

Supposez à présent que X Soit vraiment A, alors A doit répondre. La finalité du jeu, pour A, est d'essayer d'induire C en erreur. Sa réponse pourrait donc être :

« Mes cheveux sont coupés à la garçonne et les mèches les plus longues ont à peu près vingt centimètres de long. ».

Pour que le ton de la voix ne puisse pas aider l'interrogateur, les réponses devraient être écrites ou, mieux, dactylographiées. L'installation idéale serait une téléimprimante communiquant entre les deux pièces. À défaut, les questions et réponses peuvent être répétées par un intermédiaire. L'objet du jeu pour le troisième joueur (B) est d'aider l'interrogateur. La meilleure stratégie pour elle est probablement de donner les réponses vraies. Elle peut ajouter à ses réponses des choses telles que : « Je suis la femme, ne l'écoutez pas ! » mais cela ne servira à rien, car l'homme peut faire des remarques similaires.

Nous nous posons maintenant les questions: « qu'arrive-t-il si une machine prend la place de A dans le jeu ? l'interrogateur se trompera-t-il aussi souvent que lorsque le jeu se déroule entre un homme et une femme ? » Ces questions remplacent la question originale : les machines peuvent-elles penser ?

Qu'est-ce que l'intelligence, un système intelligent, une action intelligente ? Cette question, reformulée par le logicien et mathématicien A. Turing, allait dès lors propager sa quête de réponses riches dans les champs familiers des sciences humaines autant que dans ceux, plus austères, de la logique et de la mathématique, suscitant un développement médiatique sur la ligne de crête qui relie les deux versants de l'intelligence artificielle et des sciences de la cognition.

Annexe 5 : Trace d'une session Halpin-Documentaire

[M:* Bienvenue sur le système HALPIN Documentaire.

[M:0 Que souhaitez-vous faire ? 1 : Retrouver un document que vous connaissez, 2 : Rechercher un ensemble de documents, ou 3 : Découvrir ce site ?

[H:0] 2

PLAN=OUVERTURE

BUT [1=doc. particulier connu] [2=ens. de doc.] [3=Découvrir] : 0

PROFIL [1=Novice] [2=Expert] : 0

Phrase lemmatisee=2

*** Statistiques accès réseau ***

Début Traitement PERL : 10:58:18

Début accès XEROX : 10:58:18

Fin accès XEROX : 10:58:25

Début accès ELAN : 10:58:25

Fin accès ELAN : 10:58:36

Fin Traitement PERL : 10:58:36

[M:1 Rechercher un ensemble de documents. D'accord. Souhaitez-vous que je vous aide à formuler votre requête ?

[H:1] volontiers

PLAN=CHOIX_NOVICE_EXPERT

BUT [1=doc. particulier connu] [2=ens. de doc.] [3=Découvrir] : 2

PROFIL [1=Novice] [2=Expert] : 0

Phrase lemmatisee=xxAFFIRMATIONxx

Phrase lemmatisee transfo = xxAFFIRMATIONxx

Début Traitement PERL : 10:59:59

Début accès XEROX : 10:59:59

Fin accès XEROX : 11:00:05

Début accès ELAN : 11:00:05

Fin accès ELAN : 11:00:10

Fin Traitement PERL : 11:00:10

[M:2 Pouvez-vous me donner le thème recherché, entre guillemets s'il vous plaît ?

[H:2] malheureusement pas

PLAN=DETERMINER_THEME

BUT [1=doc. particulier connu] [2=ens. de doc.] [3=Découvrir] : 2

PROFIL [1=Novice] [2=Expert] : 1

Phrase lemmatisee=malheureusement xxNEGATIONxx

Phrase lemmatisee transfo = malheureusement xxNEGATIONxx

*** Statistiques accès réseau ***

Début Traitement PERL : 11:02:37

Début accès XEROX : 11:02:37

Fin accès XEROX : 11:02:43
Début accès ELAN : 11:02:44
Fin accès ELAN : 11:02:51
Fin Traitement PERL : 11:02:51

[M:3] Faites un choix parmi les critères : Auteur, Titre, Thème, Année, ou Type de document.

[H:3] je connais simplement le nom de celui qui a écrit la thèse que je cherche

PLAN=DETERMINER_CRITERE
BUT [1=doc. particulier connu] [2=ens. de doc.] [3=Découvrir] : 2
PROFIL [1=Novice] [2=Expert] : 1

Phrase lemmatisee=xxHOMMExx xxRECHERCHERxx simplement le xxAUTEURxx de celui xxQUExx avoir écrire le xxDOCUMENTxx xxQUExx xxHOMMExx xxRECHERCHERxx

Phrase lemmatisee transfo = xxHOMMExx xxRECHERCHERxx simplement yyAUTEURyy yyAMBIGUITE_Deyy celui xxQUExx yyAUTEURyy xxDOCUMENTxx xxQUExx xxHOMMExx xxRECHERCHERxx

*** Statistiques accès réseau ***
Début Traitement PERL : 11:06:24
Début accès XEROX : 11:06:24
Fin accès XEROX : 11:06:29
Début accès ELAN : 11:06:29
Fin accès ELAN : 11:06:39
Fin Traitement PERL : 11:06:39

[M:4] 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.

[H:4] cherche avec Pierrel s'il te plaît

PLAN=DETERMINER_AUTEUR
BUT [1=doc. particulier connu] [2=ens. de doc.] [3=Découvrir] : 2
PROFIL [1=Novice] [2=Expert] : 1

TYPE de la requête précédente = Thèses

Phrase lemmatisee=xxRECHERCHERxx avec Pierrel xxPOLITESSExx
Phrase lemmatisee transfo = xxRECHERCHERxx avec Pierrel yyPOLITESSEyy

2 réponses :

1. Contribution a la comprehension automatique du discours continu mise en oeuvre d'un systeme paramettable aux nivea - Pierrel, Jean-Marie - 1975
2. Etude et mise en oeuvre de contraintes linguistiques en comprehension automatique du discours continu - application aux langages artificiels - Pierrel, Jean-Marie - [s.n.] - 1981

*** Statistiques accès réseau ***
Début Traitement PERL : 11:20:34
Début accès XEROX : 11:20:34
Fin accès XEROX : 11:20:39
Début accès INRIA : 11:20:39
Fin accès INRIA : 11:20:40
Début accès ELAN : 11:20:40
Fin accès ELAN : 11:20:48
Fin Traitement PERL : 11:20:48

[M:5] J'ai trouvé 2 thèses avec comme auteur Pierrel. Consultez-les sur la page de droite. Est-ce que cela vous convient ?

[H:5] c'est exactement ce que je voulais ! merci.

[M:6] Très bien, je suis heureux d'avoir pu vous aider dans votre recherche.

PLAN=SATISFACTION

BUT [1=doc. particulier connu] [2=ens. de doc.] [3=Découvrir] : 2

PROFIL [1=Novice] [2=Expert] : 1

AUTEUR de la requête précédente = Pierrel

TYPE de la requête précédente = Thèses

Phrase lemmatisee=ce xxETRExx xxAFFIRMATIONxx ce xxQUExx xxHOMMExx xxRECHERCHERxx !
xxPOLITESSExx .

Phrase lemmatisee transfo = ce xxETRExx xxAFFIRMATIONxx xxQUExx xxHOMMExx xxRECHERCHERxx !
yyPOLITESSEyy .

*** Statistiques accès réseau ***

Début Traitement PERL : 11:28:25

Début accès XEROX : 11:28:25

Fin accès XEROX : 11:28:30

Début accès ELAN : 11:28:30

Fin accès ELAN : 11:28:35

Fin Traitement PERL : 11:28:35

Annexe 6 : Les différentes aides du système Halpin-Documentaire

Voici les différentes explications que fournit le système Halpin-Documentaire lorsque l'on sollicite son aide.

A. Aide - Accueil

Bienvenue dans l'Aide du Système HALPIN Documentaire.

HALPIN signifie : Hyperdialogue avec un Agent en Langage Proche de l'Interaction Naturelle.

Mon concepteur est José Rouillard du laboratoire CLIPS de Grenoble.
Son E-mail est Jose.Rouillard@imag.fr

Je fonctionne grâce à une Applet Java, et du code PERL.
Vous pouvez m'entendre grâce à la synthèse vocale d'Elan (www.elan.fr) .

Je travaille avec la base de données de l'INRIA (www.inria.fr)
dans le domaine de l'informatique et de l'automatique.

B. Aide sur la tâche

Je peux vous aider à effectuer une recherche documentaire (en informatique et automatique) sur le Web.
Pour cela, nous dialoguons ensemble et j'essaye de vous comprendre de mon mieux.

J'ai accès à la bibliothèque de l'INRIA (Institut National de Recherche en Informatique et en Automatique).

- ⇒ Les critères de recherche que j'utilise sont : * L'auteur ; * Le thème ; * Le titre ; * La date de publication et * Le type de document.
- ⇒ Précision :
Pour plus de précision dans un énoncé, vous pouvez utiliser des guillemets.
Exemple : le titre est "nouvelles interfaces" (sans guillemets, cette recherche ne donne pas de réponse).
- ⇒ Ajouter un élément à un critère (sauf pour la Date) :
Un élément peut être ajouté à un critère en disant "ajouter X à Y" .
Exemple : Le critère Auteur contient la valeur " Turing " et vous voulez faire une recherche sur Alan Turing. Vous pouvez dire "ajouter Alan à l'auteur".
Autres Exemples : ajouter "nouvelles interfaces" au thème. ; ajouter à l'auteur Alan s'il vous plaît.
- ⇒ Effacer (Annuler) un critère :
Un critère peut être remis à zéro en indiquant son nom. Exemple : "effacer le titre".

Un critère peut être remis à zéro en indiquant sa valeur. Exemple : "efface Boole".

Tous les critères peuvent être remis à zéro en même temps. Exemples : "annuler tout" ; "nouvelle recherche" ; "effacez tous les critères".

⇒ Concernant le critère Auteur :

Il est préférable de mettre une majuscule aux noms des auteurs.

On peut préciser un prénom ou un co-auteur (qui s'ajoute au critère auteur en cours, et qui ne l'écrase pas).

Exemples : "son prénom est Alan" ; "un des co-auteurs est Boole" .

⇒ Relancer une requête (avec les critères en cours) :

Vous pouvez soumettre une nouvelle fois la requête en cours (c'est parfois utile).

Exemples : "relance la requête" ; "refaire la recherche".

⇒ Troncature :

Vous pouvez me demander à effectuer une recherche selon des valeurs tronquées.

Exemples : "le titre commence par math" ; "le nom de l'auteur débute par Luzz".

NB : La troncature n'est possible qu'à partir de 4 caractères au moins.

⇒ Annuler la dernière opération :

Vous pouvez revenir en arrière d'un cran (c'est à dire annuler l'effet de la dernière requête).

Exemples : "annuler la dernière opération" ; "revenir en arrière" ; "défaire".

C. Aide sur la langue

Vous pouvez dialoguer avec moi en français courant.

Je travaille principalement à base de reconnaissance de concepts dans le dialogue.

Je garde en mémoire les critères que vous avez déjà utilisé et je les exploite à chaque requête.

Par exemple : Vous me dites, " l'auteur est Turing ", j'effectue une recherche sur Turing.

Si plus tard vous me dites " l'année est 1960 ", j'effectue une recherche sur Turing ET 1960.

D. Aide sur la parole

Je prononce mes réponses à travers votre navigateur grâce à une applet Java.

Si vous n'avez pas de matériel audio sur votre ordinateur, vous pouvez couper la synthèse vocale.

Exemple : "couper le son".

Les réponses seront plus rapides (car je ne charge plus le fichier son à travers le réseau).

Vous pourrez réactiver la synthèse vocale à tout moment.

Exemple : "réactiver la synthèse" .

Exemple : "remettre le son " .

E. Aide sur l'interface graphique

⇒ Bouton "Ré-écouter le son" :

Vous avez la possibilité de me faire répéter la dernière phrase prononcée en cliquant sur le bouton "Ré-écouter le son".

⇒ Augmenter/Réduire la taille de la police de caractères :

Vous pouvez augmenter ou réduire la taille de la police de caractères.

Exemple : "augmenter la taille" .

Exemple : "agrandir la police" .

Exemple : "grossir taille" .

Exemple : "réduire l'écriture" .

Exemple : "diminuer les caractères" .

Exemple : "baisser taille" .

⇒ Redimensionner les fenêtres :

Vous pouvez redimensionner toutes les fenêtres de l'interface à la souris.

Mais certaines versions de Netscape ne supportent pas cette option.

Utilisez de préférence Internet Explorer (si vous avez).

Conseil : Maximiser la fenêtre du navigateur avant de charger la page d'Halpin.

⇒ Copier/coller :

Vous pouvez copier des phrases dans la fenêtre d'historique du dialogue et les coller dans la boîte de dialogue.

⇒ Profondeur de recherche

Ce bouton est cliquable si vous avez demandé à être aidé par la machine pour formuler votre requête. Il vous permet de rechercher plus ou moins d'informations (comme les auteurs ou les thèmes proches).

Message du bouton « Profondeur de recherche » de l'applet Java :

« Ce niveau de profondeur correspond au nombre d'éléments que je récupère lors d'une recherche. Plus il est un grand, plus la recherche sera complète (mais longue). Il est valable pour l'Auteur et le thème. Exemple : en niveau 10, j'essaye de trouver 10 thèmes proches de celui recherché. »

Annexe 7 : Technique de synthèse vocale pour le Web avec Mbrola (TTS)

➤ Mbrola

Les langues disponibles avec le logiciel Mbrola standard sont :

Américain, Arabe, Portugais brésilien, Breton, Anglais, Croate, Hollandais, Estonien, Français, Allemand, Grec, Roumain, Espagnol, Espagnol mexicain, Suédois.

Il est possible de télécharger le noyau de Mbrola que l'on veut utiliser sur tel type d'environnement (20 systèmes d'exploitation sont proposés). Nous avons utilisé un noyau pour système Unix, car notre premier serveur Web était hébergé par une machine SGI.

Techniquement, dans notre application, l'appel à Mbrola se fait ainsi (serveur Unix) :

```
system ( "./mbrola ./fr1 ./$fs.pho ./$fs.wav" );
```

Dans cet exemple, \$fs.pho est le fichier de phonèmes, "fr1" est la base langagière utilisée (ici c'est le français), et \$fs.wav est le nom du fichier sonore que l'on souhaite obtenir.

Site Web du projet : <http://tcts.fpms.ac.be/synthesis/>

➤ SfConvert

Afin de convertir le fichier "wav", obtenu en sortie de Mbrola, en un fichier "au", on utilise l'utilitaire de conversion "sfconvert". On passe donc à un fichier 8 bits monophonique grâce à la commande Unix :

```
system ("sfconvert ./$fs.wav ./$fs.au -quiet rate 8000 format next");
```

➤ Encapsulation du fichier sonore dans la page Web

Cela se fait, en langage PERL, en générant le code HTML de cette manière :

```
print "<EMBED SRC=\"http://herakles.imag.fr/rouillar/dialogue/$fs.au\" AUTOSTART=true  
Hidden=true>";
```

Annexe 8 : Technique de synthèse vocale pour le Web avec Speech Engine d'Elan Informatique

Mise en pratique :

Le serveur Windows NT de notre équipe héberge le logiciel de synthèse vocale d'Elan. Un programme, développé par D. Vaufreydaz, en langage C, permet de gérer les appels au Speech Engine, provenant du Web (DosSay.exe).

Du côté serveur, voici comment se fait, en Perl, l'appel au « Speech Engine » :

```
`DosSay "$phrase_a_synthetiser" -g herakles.imag.fr -p 120 -s 160 -f $fs.au`;
```

avec comme paramètres :

-g nom_serveur (le nom du serveur où se trouve le moteur de synthèse vocale)

-p valeur_numérique (valeur du pitch, c'est-à-dire de la hauteur de son, de -600 (aigu) à 290 (grave)).

-s valeur_numérique (valeur de la vitesse d'élocution, de -600 (lent) à 290 (rapide)).

-f nom_fichier (nom du fichier que l'on souhaite obtenir au format Sun)

Du côté client, l'applet Java est chargée de récupérer le fichier son et de le jouer. Cela se code de la manière suivante :

```
fichier="synthese/" + String.valueOf(numero) + String.valueOf(compteur) + ".au";
```

```
sonajouer = getAudioClip(getCodeBase(), fichier);
```

```
sonajouer.play();
```

La variable *fichier* est calculée en concaténant le nom du répertoire cible (sur le serveur), le numéro de la machine cliente, un compteur de session et l'extension au format Sun du fichier audio. Ensuite, on effectue une requête sur le Web, et l'on récupère dans la variable *sonajouer*, le fichier correspondant à la synthèse vocale voulue, puis on le fait jouer par la biais du navigateur.

Annexe 9 : Exemples de DHM recueillis avec Halpin-Documentaire

Ci-dessous, quelques exemples de dialogues où l'on voit que la machine adapte ses interventions grâce au contexte (nombre de documents retrouvés, rappel des critères de recherches, relance, etc.)

14:26:44:M7=J'ai trouvé 7 documents de 1998 ayant comme thème parole . Consultez-les sur la page de droite. Est-ce que cela vous convient ?

14:29:48:H7=oui je veux ceux de 99

14:30:02:M8=Très bien, je suis heureux d'avoir pu vous aider dans votre recherche.

14:30:27:H8=je veux ceux de 99 aussi

14:30:36:M9=Est-ce que 1999 est bien la date de publication.

14:30:54:H9=oui

14:31:08:M10=J'ai trouvé 2 documents de 1999 ayant comme thème parole. Consultez-les sur la page de droite. Est-ce que cela vous convient ?

14:31:43:H10=oui fin

* * *

14:24:13:M8=Connaissez-vous le nom de l'auteur du document que vous recherchez ?

14:24:49:H8=Hawkins

14:25:02:M9=J'ai trouvé un document avec comme auteur Hawkins. Consultez-le sur la page de droite. Est-ce bien celui que vous cherchiez ?

* * *

16:40:37:M0=Que souhaitez-vous faire ? 1 : Retrouver un document que vous connaissez, 2 : Rechercher un ensemble de documents, ou 3 : Découvrir ce site ?

16:40:37:H0=1

16:40:59:M1=Retrouver un document que vous connaissez. D'accord. Souhaitez-vous que je vous aide à formuler votre requête ?

16:41:17:H1=non, l'auteur est Caelen

16:41:30:M2=Très bien, je vous laisse faire. Appelez à l'aide si vous êtes en difficulté.

16:41:46:H2=l'auteur est Caelen

16:41:55:M3=J'ai trouvé 2 documents avec comme auteur Caelen.

16:42:31:H3=Bene, maintenant le theme est "mathématiques"

16:44:10:M4=Je n'ai pas trouvé de document avec comme auteur Caelen ayant comme thème mathématiques.

* * *

21:24:50:H25=je cherche des documents sur l'histoire du vingtième siècle.
21:25:02:M26=Je prépare une requête ayant comme thème - vingtième siècle - . Nous sommes bien d'accord ?
21:25:41:H26=oui cela a du sens.
21:25:54:M27=Je n'ai pas trouvé de document ayant comme thème vingtième siècle⁵³ .
21:26:40:H27=Cherchez à histoire

* * *

10:40:10:M7=J'ai trouvé 8 documents ayant comme thème virus.
10:40:46:H7="moniteur"
10:40:56:M8=Est-ce que moniteur est bien l'auteur du document.
10:41:29:H8=non
10:41:37:M9=J'ai trouvé 42 documents ayant comme thème moniteur.
10:41:56:H9=as tu du feu ?
10:42:08:M10=Je prépare une requête ayant comme thème - feu - . Nous sommes bien d'accord ?
10:42:57:H10=non
10:43:01:M11=Alors, reformulez votre phrase s'il vous plaît.

* * *

17:48:22:H2=au secours
17:48:32:M3=Okay voici une fenêtre d'aide. Souhaitez-vous que je vous aide à formuler votre requête ?
17:57:15:H3=oui
17:57:20:M4=Pouvez-vous me donner le thème recherché, entre guillemets s'il vous plaît ?

* * *

16:24:20:M1=Retrouver un document que vous connaissez. D'accord. Souhaitez-vous que je vous aide à formuler votre requête ?
16:24:33:H1=non
16:24:39:M2=Très bien, je vous laisse faire. Appelez à l'aide si vous êtes en difficulté.
16:25:10:H2=je cherche un document qui s'intéresse à la comparaison entre Linux et Windows NT
16:25:20:M3=Donnez-moi plus d'indications à propos de ce document.
16:25:41:H3=comparer windows NT et Linux
16:25:51:M4=windows NT Linux sont des mots que je ne connais pas.

* * *

⁵³ L'ordinateur n'a pas rectifié l'accentuation du mot siècle, car il n'y a pas eu de lemmatisation sur « vingtième siècle » (cependant, la BD de l'INRIA ne possède pas de documents répondant à cette requête, même bien orthographiée).

Annexe 10 : Questionnaire de satisfaction Halpin-Documentaire

Merci de donner votre avis à propos du système Halpin-Documentaire :

LE DIALOGUE HOMME-MACHINE :

Globalement, l'interaction dialoguée avec le système HALPIN est :

- Très mauvaise
- Mauvaise
- Bonne
- Très bonne

Le dialogue avec la machine est efficace :

- Pas du tout
- Un peu
- Assez
- Très

La machine comprend bien ce que vous voulez :

- Jamais
- De temps en temps
- La plupart du temps
- Toujours

La coopération de la machine :

- Insuffisante
- Correcte
- Trop lourde

Le dialogue avec la machine vous a-t-il orienté dans votre recherche :

- Pas assez
- Un peu
- Suffisamment
- Beaucoup

Le dialogue vous a permis de trouver ce que vous cherchiez :

- Pas du tout
- Un peu
- En partie
- Totalelement

A propos du dialogue en langue naturelle, vous trouvez le système :

- Pas du tout robuste
- Peu robuste
- Robuste
- Très robuste

Vous avez apprécié l'aide de la machine :

- Oui
 - Non
-

LA SYNTHÈSE VOCALE :

Globalement, elle est :

- Sans opinion
- Très mauvaise
- Mauvaise
- Bonne
- Très bonne

Vous l'avez utilisée :

- Sans opinion
- Jamais
- Une ou deux fois pour essayer
- Presque toujours
- Pendant tout le dialogue

Vous l'avez interrompue :

- Sans opinion
- Jamais
- Une ou deux fois pour essayer
- Presque toujours
- A chaque fois

Vous l'avez réécoutée :

- Sans opinion
- Jamais
- Une ou deux fois pour essayer
- Presque toujours
- A chaque fois

Elle est intelligible :

- Sans opinion
- Très peu
- Peu
- Moyennement
- Très

Elle semble naturelle :

- Sans opinion
- Très peu
- Peu
- Moyennement
- Très

Elle améliore l'interaction :

- Sans opinion
- Très peu
- Peu
- Assez
- Beaucoup

LA RECONNAISSANCE VOCALE :

Globalement, elle est :

- Sans opinion
- Très mauvaise
- Mauvaise
- Bonne
- Très bonne

L'avez-vous utilisé :

- Sans opinion
- Jamais
- Une ou deux fois pour essayer
- Presque toujours
- Pendant tout le dialogue

Elle améliore l'interaction :

- Sans opinion
- Très peu
- Peu
- Assez
- Beaucoup

Elle permet de s'exprimer :

- Sans opinion
- Moins bien qu'avec le clavier
- Aussi bien qu'avec le clavier
- Plus qu'avec le clavier

Elle permet de dire :

- Sans opinion
- Moins de choses qu'avec le clavier
- Autant de choses qu'avec le clavier
- Plus de choses qu'avec le clavier

Avez-vous alterné entre interaction clavier/souris et interaction vocale :

- Sans opinion
- Jamais
- De temps en temps
- Souvent
- Toujours

UTILITE DU SYSTEME :

Pour une recherche documentaire, par rapport à un formulaire traditionnel du Web, diriez-vous que le système HALPIN est :

- Largement moins bien
- Moins bien
- Pareil
- Mieux
- Largement mieux

Vous avez trouvé le dialogue avec la machine :

- Trop court
- Normal
- Trop long

Le temps de réponse du système est :

- Très rapide
- Rapide
- Acceptable
- Trop longtemps

VOUS :

Vous êtes :

- Un homme

- Une femme
- Une machine

Vous êtes :

- Etudiant(e)
- Enseignant et/ou Chercheur
- Salarié(e)
- Autre

Vous êtes informaticien(ne) :

- Oui
- Non

Vous êtes bibliothécaire :

- Oui
- Non

En informatique, vous êtes :

- Novice
- Confirmé
- Expert

En recherche d'information, vous êtes :

- Novice
- Confirmé
- Expert

Vous avez demandé à être guidé par la machine :

- Oui (toujours)
- Non (jamais)
- Oui et Non (selon les moments)

COMMENTAIRE LIBRE : -----

Annexe 11 : Le projet ORION

L'une des voies que ce projet propose d'explorer est la conception de cartes cognitives multimodales pour le Web. Cette 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'utilisateurs bien déterminée et (b) une réflexion sur les modes d'interactions (vision, parole, geste ...) qui pourraient faciliter cette activité. Il apparaît en effet essentiel, avant de proposer toute nouvelle interface (multimodale ou pas), de préciser quels sont les freins et/ou les motivations que les utilisateurs d'Internet rencontrent actuellement.

Nous nous sommes également penché sur la prise en compte des besoins et des profils des utilisateurs. Toutefois, pour adapter les navigateurs actuels à chaque type d'utilisateur, encore faut-il connaître les caractéristiques de chacun d'entre eux (les utilisateurs) [BEAUFILS 96]. Il convient alors de déterminer comment sont organisées les connaissances des utilisateurs du Web (d'une manière statique) et comment ils les utilisent dans un contexte précis comme celui de la recherche d'information dans un environnement hypertextuel (de manière dynamique, donc).

Par ailleurs, nous nous sommes inspiré de travaux qui tentent de rendre le Web moins impersonnel (notamment [EKLUND et al. 96]). C'est-à-dire que l'outil de navigation s'adapte à son utilisateur en observant ses comportements et ses habitudes, puis tente d'anticiper ses actions de manière intelligente au cours de la navigation, par exemple en chargeant à l'avance les pages d'un site susceptibles d'intéresser l'utilisateur où en se focalisant sur des informations proches de son profil [BOUHAI et al. 99].

Les résultats obtenus dans le cadre du projet Orion sont les suivants :

❖ **Projet Orion : l'aspect visuel, et le navigateur Nestor**

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.

Un processus de conception commence généralement par une étude d'usages. Mais les utilisateurs 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.

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 (cf. Figure 38). 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'utilisateur navigue, avec une mise en page axée sur l'idée de "trajet". L'utilisateur-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 sauveées 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 ses propres réactions 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").

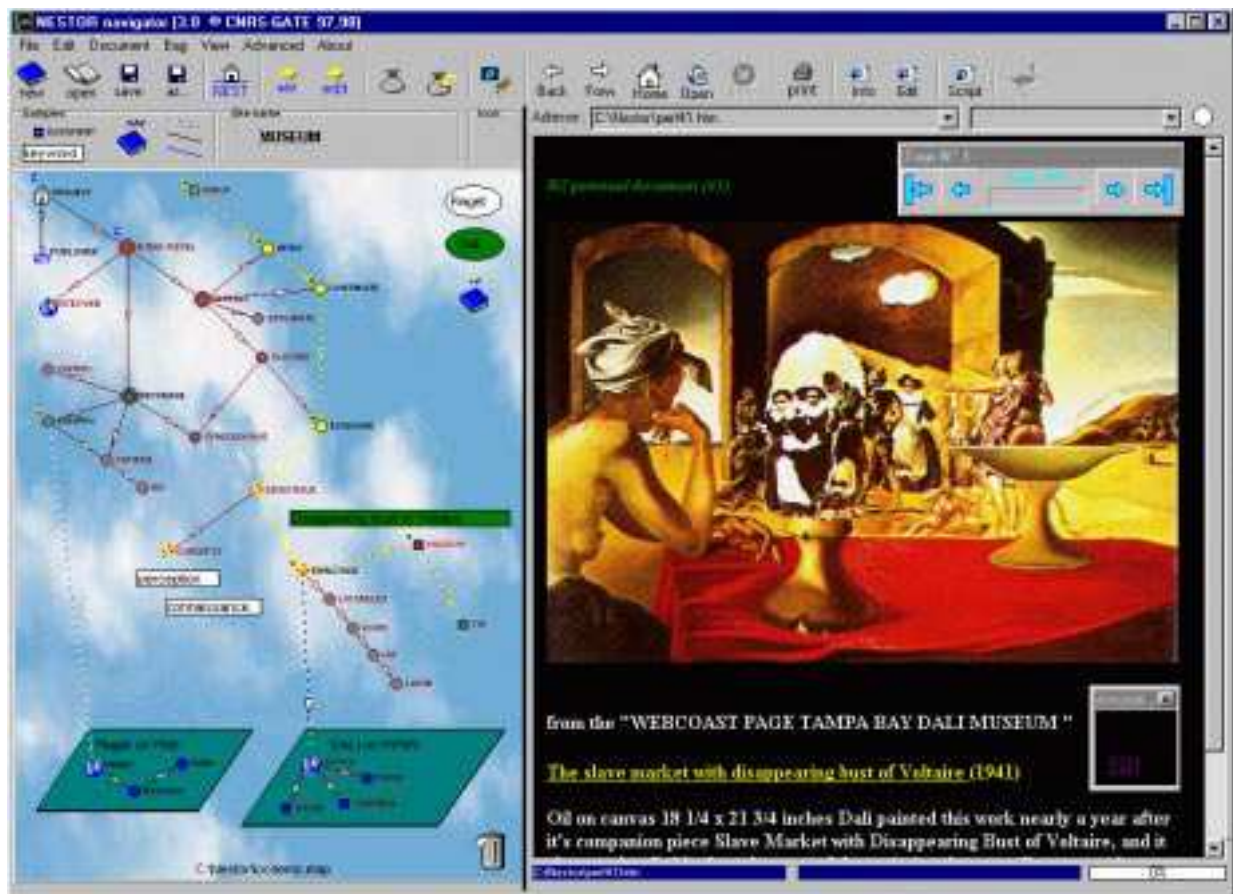


Figure 38 : Une vue générale de l'écran de NESTOR

La Figure 38 ci-dessus présente une vue générale de l'écran de NESTOR (à gauche) intégré avec InternetExplorer (à droite). On notera la présence dans la carte de gauche, de deux « zones conceptuelles », de nombreux documents et liens personnels (en pointillés), et d'un « parcours guidé » (en pointillés) à partir duquel est généré le menu de navigation linéaire simplifié présent en haut de l'image de droite. Le domaine abordé ici est celui des « illusions d'optique » dans le cadre d'un cours sur la psycho-pédagogie et l'audiovisuel (STE Université de Liège, Belgique).

La version finale⁵⁴ intègre les résultats des deux pôles puisque l'on peut utiliser le système HALPIN au sein du navigateur NESTOR. D'autres navigateurs de ce type et proches du principe de NESTOR sont présentés succinctement en annexe 1.

❖ **Projet Orion : l'aspect oral, et le système Halpin**

Le pôle grenoblois a choisi de travailler sur l'aspect oral dans le domaine de la recherche documentaire. En effet, 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

⁵⁴ Téléchargeable à l'adresse <http://www.gate.cnrs.fr/~zeiliger/nestor.htm>

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'utilisateur. Dans ce sens, des travaux montrent comment l'information peut être présentée de manière originale à l'utilisateur : Vernier et Nigay proposent une représentation d'une grande quantité d'information avec leur logiciel *Vitesse*⁵⁵, fondée sur une perception en œil de poisson (fish eye view) [VERNIER et NIGAY 97]. Mais cela reste une approche visuelle du problème.

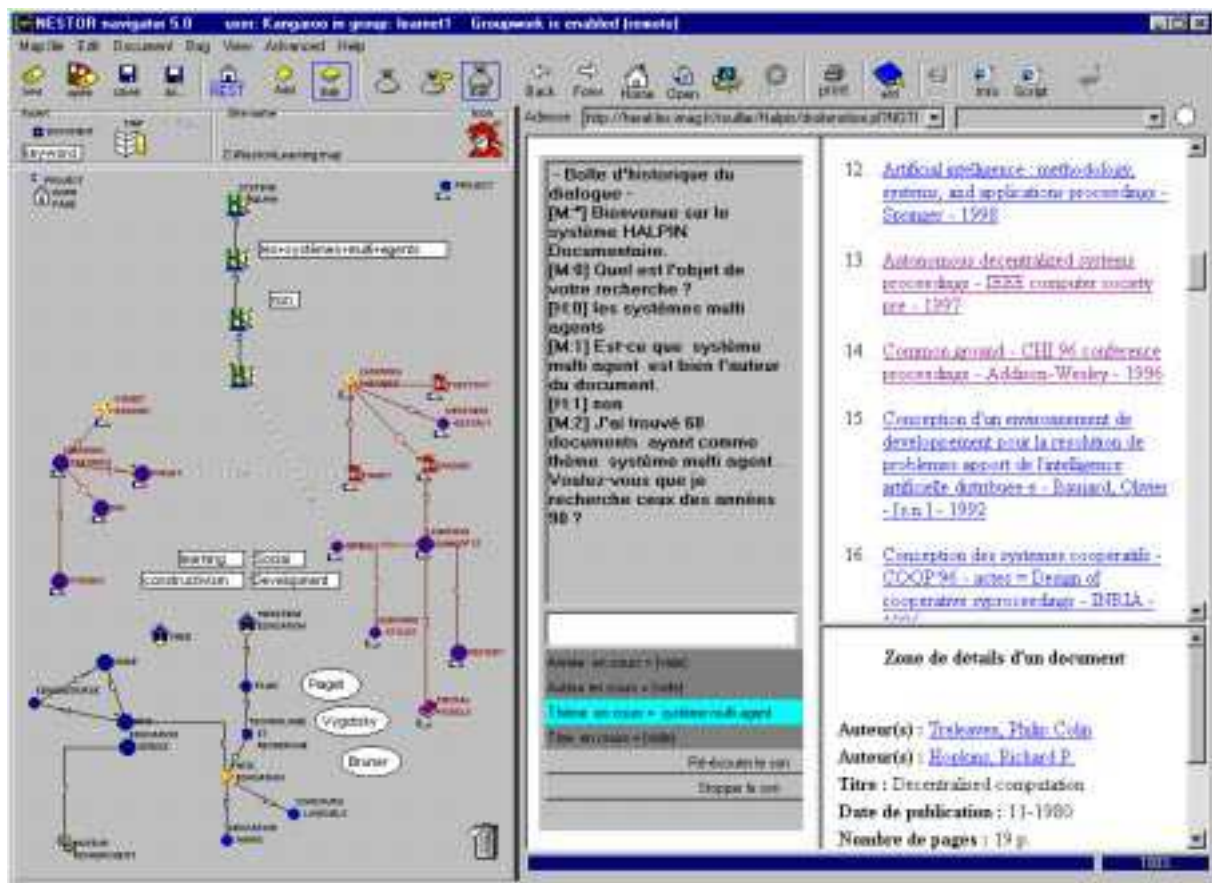
L'approche du groupe grenoblois consiste, non pas à écarter ces canaux de communications, mais à proposer en sur-couche aux navigateurs actuels, des dispositifs auditifs d'entrée et de sortie de données, permettant une communication et un dialogue oral, en LN, avec l'utilisateur.

Comme pour le pôle lyonnais, des enquêtes d'usages 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 préalables nous ont permis d'étudier 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.

Le projet Orion a servi de cadre méthodologique pour cette thèse. Il prévoyait théoriquement de passer par les phases suivantes : observation et analyse d'utilisateurs d'Internet dans le cadre d'activités de recherche d'informations, étude d'une architecture générique pour l'intégration d'une couche de dialogue, étude et réalisation du système de dialogue, démonstration et validation du système ; puis évaluation dans le domaine choisi par le projet Orion (éducation et formation).

Les deux systèmes développés dans le cadre de ce projet sont utilisables simultanément. La Capture 24 ci-après montre l'interface du navigateur NESTOR au cours d'une recherche d'information grâce au système HALPIN : à droite, un dialogue à la fois textuel et sonore guide l'utilisateur dans sa recherche de document, et à gauche, une « carte » de l'espace hypermédia visité a été tracée automatiquement ; elle a été enrichie et réorganisée par l'utilisateur.

⁵⁵ <http://iihm.imag.fr/demos/VitesseDemos.html>



Capture 24 : Interface du navigateur NESTOR lors d'une utilisation du système HALPIN

Comme nous pouvons le voir sur la Figure 40, ci-après, 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 :

Dialogue (A,B) ➔ requête (A,B) + promesse (B, A) + information (B,A) + test-satisfaction (A,B)

ou bien

Dialogue (A,B) ➔ offre (B,A) + acceptation (A, B) + information (B,A) + test-satisfaction (A,B)

La Figure 40 représente un réseau récursif de transitions dans un dialogue : les cercles sont des états dans le dialogue, les flèches représentent des transitions d'un état vers un autre (correspondant aux actes de dialogue *requête*, *offre*, *promesse*, *acceptation*, *information* et *évaluation*), et les carrés sont des états finaux.

Dans le premier cas (états 1,2,3,4), le dialogue entre A et B commence par une requête de A à B. Le paramètre T indique le type de dialogue (recherche d'information ou métadialogue). Le dialogue se poursuit avec la promesse (de B à A) que l'information va être fournie. Dans un troisième temps, après une éventuelle interrogation de la base de données, B fournit l'information promise à A. Puis A évalue la pertinence de l'information présentée. Le dialogue peut alors se terminer (état final n°5), ou bien se poursuivre (en bouclant vers le nœud n°1).

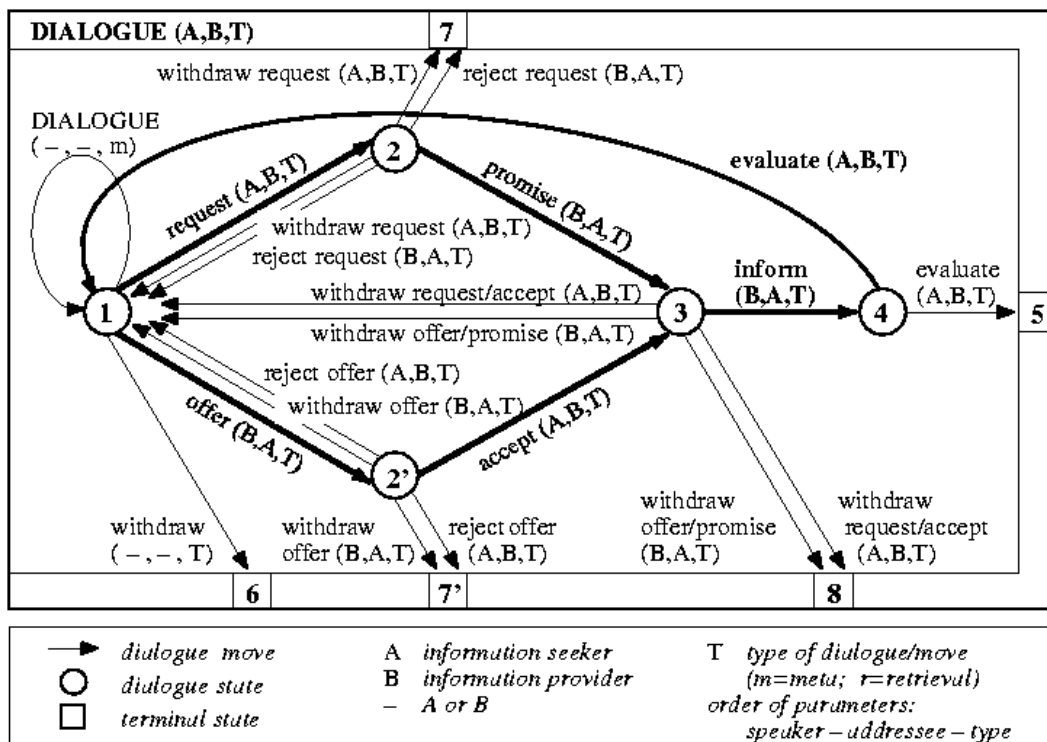


Figure 40 : Le modèle COR (Conversational Roles)

Dans le second cas (1,2',3,4), le dialogue entre A et B commence à l'initiative de B, qui propose d'offrir de l'information à A. Ce dernier peut accepter ou rejeter cette proposition. S'il accepte, B lui fournit cette information, et la suite du dialogue est similaire au premier cas évoqué ci-dessus.

Avec ce réseau récursif de transitions dans un dialogue, une transition (une flèche sur le schéma) n'est pas considérée comme un élément atomique, mais comme un sous-réseau de transitions. La Figure 41 ci-après, montre le sous-réseau de l'offre d'information entre A et B.

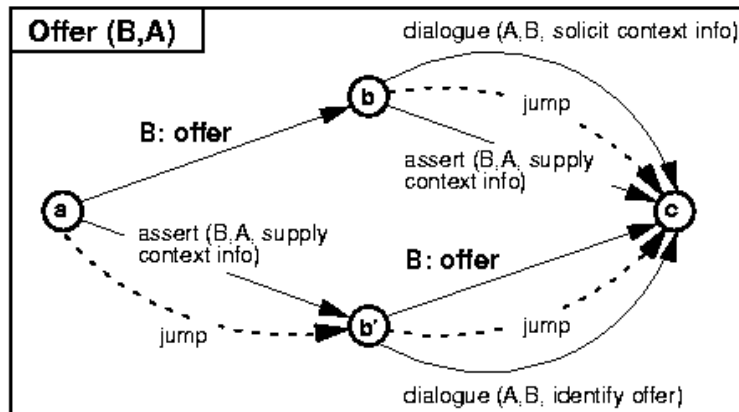


Figure 41 : Sous-réseau détaillé de l'offre (B,A) dans le modèle COR

En effet, dans un dialogue, lorsque B propose d'offrir une information à A, il se peut que les interlocuteurs ne se comprennent pas immédiatement, ou que le dialogue soit ambigu. L'offre (B,A) est donc vue comme un sous-dialogue, qui peut prendre plusieurs tours de parole.

Pour plus d'informations, voir la page Web de Delite (Advanced retrieval support for digital libraries), consulter :

<http://delite.darmstadt.gmd.de/delite/Publications/>

Annexe 13 : Bibliographie personnelle relative à ces travaux de thèse

Les références ci-dessous ne sont pas citées dans la bibliographie générale de la thèse

Documents publiés :

[CAELEN et ROUILLARD 99] CAELEN, J., ROUILLARD, J., *Le système HALPIN : recherche documentaire en langue naturelle*. Workshop NTE-SPI-Sante, Lyon, 1999.

[CAELEN et ROUILLARD 98] CAELEN, J., ROUILLARD, J., *Interface vocale pour la recherche documentaire*. Workshop NTE-SPI-Sante, Lyon, 1998.

[ROUILLARD 99b] 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, 1999, pp. 3-20.

[ROUILLARD 99c] ROUILLARD, J., *HALPIN : A natural language information retrieval system for a digital library on the World Wide Web*, Webnet'99, Waikiki Beach, Honolulu, Hawaii, USA, 1999.

[ROUILLARD 99d] ROUILLARD, J., *Navigation versus dialogue sur le Web. Une étude des préférences.IHM'99*, Montpellier, 1999, pp 53-56.

[ROUILLARD 98b] ROUILLARD, J., *Contribution à l'étude du dialogue Homme-Machine à travers le Web : la personnalisation*, RECITAL'98, Le Mans, 1998, pp 127-134.

[ROUILLARD 97] ROUILLARD, J., *Vers un dialogue oral sur le World Wide Web*, 2èmes Rencontres Jeunes Chercheurs en Parole, La Rochelle, 1997.

[ROUILLARD et CAELEN 99] ROUILLARD, J., CAELEN, J., *Multimodal Information Seeking Dialogues on the World Wide Web*, Eurospeech'99, Budapest, Hungary, 1999.

[ROUILLARD et CAELEN 99b] ROUILLARD, J., CAELEN, J., *A Multimodal and conversational application in natural language for information seeking on the World Wide Web : The HALPIN system*, 4th International Conference on Applications of Natural Language to Information Systems , University of Klagenfurt, Austria, 1999, pp 239-244.

[ROUILLARD et CAELEN 99c] ROUILLARD, J., CAELEN, J., *HALPIN : a multimodal and conversational system for information seeking on the World Wide Web*, ESCA ETRW workshop : Accessing information in spoken audio, Cambridge (UK), 1999, pp 20-24.

[ROUILLARD et CAELEN 97] ROUILLARD, J., CAELEN, J., *Etude de la propagation au sein du Web à travers les liens hypertextes*, 4ème conférence Internationale Hypertextes et Hypermédiâs - Septembre 1997, Paris. Numéro spécial de la revue Hypertextes et Hypermédiâs, éditions Hermès, 1997, Paris, pp 87-100.

[ROUILLARD et CAELEN 97b] ROUILLARD, J., 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, 1997, Seoul, Korea.

[VAUFREYDAZ et al. 98] VAUFREYDAZ, D., MEIYE, J-P., ROUILLARD, J., *Webcognition, expérimentations cognitives sur Internet*, 6ème journée d'étude de l'association Accion , CREPCO, Aix en Provence, 1998.

[VAUFREYDAZ et al. 99] VAUFREYDAZ, D., ROUILLARD, J., AKBAR, M., *A network architecture for building applications that use speech recognition and/or synthesis*, Eurospeech'99, Budapest, Hungary, 1999.

[VAUFREYDAZ et al. 99b] VAUFREYDAZ, D., AKBAR, M., ROUILLARD, J., *Internet Documents: A Rich Source for Spoken Language Modeling*, Automatic Speech Recognition and Understanding (ASRU'99), IEEE, Keystone, Colorado, USA, 1999.

Documents non publiés :

[ROUILLARD 96] ROUILLARD, J. *Une composante orale de dialogue pour Internet. FACOPI : Feuilleter A Commandes Orales Pour Internet*. Université Joseph Fourier de Grenoble. Rapport de DEA en Informatique Systèmes et Communications, Grenoble, 1996

[ROUILLARD 98c] ROUILLARD, J., *Dialogue oral Personne-Système sur le Web*, Contribution à l'atelier thématique "Interaction sur le WWW" de la conférence IHM'98, Nantes, 98.

[ROUILLARD 99e] ROUILLARD, J., *Rapport d'activité scientifique à la région Rhône-Alpes pour le projet ORION (3ème année). Une composante orale de dialogue pour Internet*. Grenoble, Avril 1999.

[ROUILLARD 98d] ROUILLARD, J., *Rapport d'activité scientifique à la région Rhône-Alpes pour le projet ORION (2ème année). Une composante orale de dialogue pour Internet*. Grenoble, Avril 1998.

[ROUILLARD 97b] ROUILLARD, J., *Rapport d'activité scientifique à la région Rhône-Alpes pour le projet ORION (1ère année). Une composante orale de dialogue pour Internet*. Grenoble, Avril 1997.

[ROUILLARD 97c] ROUILLARD, J., *Enquête d'usage à la médiathèque IMAG*, Rapport dans le cadre du projet ORION. Grenoble, Janvier 1997.

[ROUILLARD 99f] ROUILLARD, J., *Présentation du système Halpin Documentaire au salon TEC'99 (Carrefour européen des technologies et de la compétitivité)*, Grenoble, 1999.

[CAELEN et al. 99] CAELEN, J., AKBAR, M., ROUILLARD, J. Prospective sur les interfaces vocales. Rapport final de contrat EDF, Décembre 99.

Table des matières

1	INTRODUCTION.....	6
1.1	LE DIALOGUE HOMME-MACHINE	6
1.2	LE RÉSEAU INTERNET	9
1.3	MOTIVATION DE NOS TRAVAUX : LE PROJET ORION	12
1.4	MÉTHODOLOGIE.....	13
2	SITUATION DU PROBLÈME : DIALOGUE ET RECHERCHE D'INFORMATION	16
2.1	LES SYSTÈMES DE DHM	16
2.2	LA RECHERCHE D'INFORMATION SUR INTERNET	38
2.2.1	<i>Recherche géographique.....</i>	<i>39</i>
2.2.2	<i>Recherche thématique</i>	<i>39</i>
2.2.3	<i>Recherche par mots-clés.....</i>	<i>40</i>
3	BESOINS DES UTILISATEURS EN SITUATION DE RECHERCHE D'INFORMATION.....	43
3.1	CONNAISSANCES POUR LE MODÈLE DE LA TÂCHE.....	45
3.2	CONNAISSANCES POUR LE MODÈLE DE L'UTILISATEUR.....	47
3.3	CONNAISSANCES POUR LE MODÈLE DE DIALOGUE	49
4	MISE EN ŒUVRE D'UN SYSTÈME DE DHM	53
4.1	PROBLÈME DE CIRCULARITÉ POUR LA MISE EN ŒUVRE D'UN DHM.....	54
4.2	CRITIQUE DE LA MÉTHODE MAGICIEN D'OZ.....	55
4.3	CRITIQUE DE LA MÉTHODE MESSAGERIE ÉLECTRONIQUE.....	55
4.4	SYSTÈME GÉNÉRIQUE D'HYPERDIALOGUE.....	57
4.5	APPORTS DE L'HYPERDIALOGUE	58
4.5.1	<i>Apports aux systèmes de DHM.....</i>	<i>59</i>
4.5.2	<i>Apports à l'étude du DHM</i>	<i>59</i>
5	LE SYSTÈME HALPIN-RECUEIL	61
5.1	PRINCIPE DE L'EXPÉRIENCE DE RECUEIL DE DIALOGUES.....	62
5.1.1	<i>Intérêt de cette méthode.....</i>	<i>62</i>
5.1.2	<i>Modalités techniques.....</i>	<i>63</i>
5.2	ARCHITECTURE DU SYSTÈME HALPIN-RECUEIL	64
5.3	FONCTIONNEMENT DU SYSTÈME D'INTERACTION.....	66
5.4	RÉSULTATS POUR LE SYSTÈME HALPIN-RECUEIL	70
5.4.1	<i>Présentation des résultats.....</i>	<i>70</i>

5.4.1.1	Exemples de données recueillies avec le système Halpin-Recueil	71
5.4.1.2	Les tours de parole	74
5.4.1.3	Les fréquences lexicales	76
5.4.2	<i>Observations similaires aux corpus de DHM classiques</i>	79
5.4.2.1	Le dialogue : plus qu'un échange de questions et de réponses.....	79
5.4.2.2	Co-construction d'un vocabulaire commun : l'homme s'adapte à la machine et vice-versa	82
5.4.2.3	Statistiques sur l'emploi de termes spécifiques à la tâche	84
5.4.2.4	Les verbes utilisés par les utilisateurs	86
5.4.2.5	Les substantifs utilisés par les utilisateurs.....	87
5.4.3	<i>Observations différentes par rapport aux corpus de DHM classiques</i>	88
5.4.3.1	La mise en doute des informations données par la machine.....	90
5.4.3.2	La personnalisation	91
5.4.3.3	Influence du support (Internet) sur le dialogue	94
6	LE SYSTÈME HALPIN-DOCUMENTAIRE.....	99
6.1	PRÉSENTATION DU SYSTÈME.....	99
6.2	RECONNAISSANCE DE CONCEPTS	102
6.2.1	<i>Principe</i>	102
6.2.2	<i>Des concepts aux actes</i>	106
6.3	MODÈLE DE DIALOGUE.....	106
6.3.1	<i>Gestion du dialogue</i>	109
6.3.1.1	Gérer les tours de parole.....	110
6.3.1.2	Calculer les interventions du système	112
6.3.1.3	Prédire les interventions de l'utilisateur.....	114
6.3.1.4	Interpréter les interventions de l'utilisateur.....	119
6.3.1.5	Adapter la stratégie de dialogue à la situation	120
6.4	ARCHITECTURE DU SYSTÈME	125
6.5	COMPOSANTES ORALES DE DHM.....	127
6.5.1	<i>La synthèse vocale</i>	128
6.5.2	<i>La reconnaissance de la parole</i>	131
6.5.2.1	Reconnaissance locale.....	132
6.5.2.2	Reconnaissance distante.....	135
6.6	INTERFACE MULTIMODALE DU SYSTÈME	136
7	RÉSULTATS ET ÉVALUATIONS DU SYSTÈME HALPIN-DOCUMENTAIRE.....	141
7.1	PRÉSENTATION DES RÉSULTATS	141
7.2	EVALUATION DU SYSTÈME	143
7.2.1	<i>Exemples de DHM avec le système Halpin-Documentaire</i>	144
7.2.2	<i>Critique du système Halpin-Documentaire par les utilisateurs</i>	148
7.2.3	<i>Evaluation du temps de réponse de la machine</i>	150
7.2.4	<i>Intégration d'une vue en œil de poisson</i>	151
7.2.4.1	Dialogue versus Navigation	153
7.2.4.2	Réponses au sondage concernant la FEV	155
7.2.4.3	Conclusion sur l'intégration de la FEV dans Halpin-Documentaire	157

7.3	LIMITES DU SYSTÈME HALPIN-DOCUMENTAIRE	158
7.3.1	<i>Limites techniques</i>	158
7.3.2	<i>Limites dialogiques</i>	159
8	CONCLUSION ET PERSPECTIVES	163
	BIBLIOGRAPHIE	169
	TABLE DES ILLUSTRATIONS	181
	ANNEXES.....	184
	ANNEXE 1 : QUELQUES OUTILS DE NAVIGATION CARTOGRAPHIQUE POUR LE WEB.....	184
	ANNEXE 2 : LE QUESTIONNAIRE DE L'ENQUÊTE D'USAGE À LA MÉDIATHÈQUE IMAG.....	188
	ANNEXE 3 : UTILISATION D'INTERFACES DE DHM ORAL DANS LE DOMAINE MILITAIRE	191
	ANNEXE 4 : LE TEST DE TURING.....	193
	ANNEXE 5 : TRACE D'UNE SESSION HALPIN-DOCUMENTAIRE	194
	ANNEXE 6 : LES DIFFÉRENTES AIDES DU SYSTÈME HALPIN-DOCUMENTAIRE	197
	ANNEXE 7 : TECHNIQUE DE SYNTHÈSE VOCALE POUR LE WEB AVEC MBROLA (TTS)	200
	ANNEXE 8 : TECHNIQUE DE SYNTHÈSE VOCALE POUR LE WEB AVEC SPEECH ENGINE D'ELAN INFORMATIQUE.....	201
	ANNEXE 9 : EXEMPLES DE DHM RECUEILLIS AVEC HALPIN-DOCUMENTAIRE	202
	ANNEXE 10 : QUESTIONNAIRE DE SATISFACTION HALPIN-DOCUMENTAIRE.....	204
	ANNEXE 11 : LE PROJET ORION.....	208
	ANNEXE 12 : LE MODÈLE COR	213
	ANNEXE 13 : BIBLIOGRAPHIE PERSONNELLE RELATIVE À CES TRAVAUX DE THÈSE	216