



HAL
open science

L'hyperdialogue

José Rouillard

► **To cite this version:**

José Rouillard. L'hyperdialogue. Céline Paganelli. Interfaces Homme-Machine et Recherche d'information, Hermès-Lavoisier, pp 229-259, 2002, 9782746204263. hal-02438974

HAL Id: hal-02438974

<https://hal.science/hal-02438974>

Submitted on 14 Jan 2020

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.

Chapitre 7

L'hyperdialogue

Depuis bien des années, l'information semble prédominer sur tous les autres éléments de notre monde. S'il est aisé de se rendre compte que les moyens de transmission de cette information évoluent, il est en revanche plus subtil de percevoir que petit à petit, les agents qui la transmettent évoluent, eux aussi. Les hommes n'utilisent plus seulement des machines pour une communication médiatisée (le téléphone mettant en relation deux interlocuteurs humains, par exemple), mais également pour interagir directement avec un agent artificiel. L'homme dialogue donc avec une machine pour retrouver de l'information. Si ce phénomène n'est pas nouveau (l'Eliza de Weizenbaum date des années 1960, [WEI 66]), ce qui l'est plus, c'est que ces machines *commencent* à comprendre l'homme.

Le dialogue sur le réseau internet nous intéresse particulièrement, et nous avons proposé une modélisation du dialogue homme-machine en langue naturelle, sur le web, grâce à la notion d'hyperdialogue. Dans ce chapitre, nous présentons ce concept ainsi que le système Halpin-Documentaire. Cette réalisation informatique donne une idée relativement précise de ce que la machine peut ou ne peut pas faire, ce qu'elle croit comprendre et ce qu'elle comprend réellement des interventions de l'utilisateur en situation de recherche d'informations sur le web. Nous présentons ci-après le modèle de dialogue conçu, l'architecture logicielle mise en œuvre, le système complet réalisé (avec reconnaissance et synthèse vocale) ainsi que des exemples réels de dialogues homme-machine recueillis au cours des différentes campagnes d'évaluation de ce système de recherche d'informations documentaires.

7.1. A quoi servent les ordinateurs ?

A bien y réfléchir, la réponse à la question « à quoi servent les ordinateurs ? » n'est pas si évidente qu'il n'y paraît. Certes, la puissance de calcul sans cesse croissante de nos machines est indiscutable, et il est évident qu'elles surpassent la plupart d'entre nous, lorsqu'il s'agit de compter. Mais est-ce tout ce qu'on leur demande ? En schématisant, on pourrait dire que l'ordinateur est capable de réaliser une tâche, à condition qu'on lui fournisse les bonnes entrées. Or, on ne donne pas des informations à un ordinateur de la même manière qu'on le ferait avec un interlocuteur humain (ou du moins, pas encore). Il faut, en effet, lors d'une interaction avec une machine, respecter des règles, des formats, des structures d'échange de données, et autres protocoles de communication. En d'autres termes, il faut quelquefois que l'utilisateur connaisse et transmette non seulement l'information elle-même, mais également la nature de cette dite information.

Si les ordinateurs nous sont d'une grande utilité (car rapides et précis) lorsque l'on sait effectivement ce que l'on recherche, en revanche, ils nous paraissent bien peu appropriés pour nous aider dès lors que nous ne pouvons formuler pleinement les éléments qui composent l'objet de cette recherche d'information (RI). En fait, on considère généralement qu'une recherche d'information devient efficace lorsqu'elle met en jeu un processus itératif et de raffinement qui permet de se rapprocher, pas à pas, de l'information convoitée. Cela est particulièrement vrai sur le réseau internet :

« Satisfying an information need through on-line full-text searching is often an iterative process in which the user successively modifies an initial query in response to the results obtained » [ANI 98].

Et c'est notamment sur cet aspect que réside une grande partie des travaux de recherche en communication homme-machine. L'homme n'est certes pas toujours en mesure d'apporter la bonne information au bon moment ; il est cependant, la plupart du temps, capable de décider si tel renseignement qu'on lui fournirait serait éventuellement intéressant pour faire avancer le processus global de recherche. Il faut alors que les interlocuteurs se comprennent, et que tour à tour ils argumentent, s'expliquent, se convainquent. Nous faisons cela naturellement en dialoguant avec un libraire ou une bibliothécaire. Ne pourrions-nous pas dialoguer de la même manière avec une machine ? Serait-ce utile et surtout efficace ?

Il est vrai que 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 » [NIC 99].

Nous sommes passés depuis quelques années de systèmes technocentrés à des systèmes beaucoup plus anthropocentrés, pour lesquels l'utilisateur est au cœur du processus d'interaction. Les chercheurs du domaine tentent donc d'amener la machine à comprendre la langue naturelle (LN) : « Exprimer son problème en langage naturel à une machine, n'est-ce pas le but ultime des recherches en programmation ? » [COU 86]. 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 1960, de nombreux systèmes informatiques ont été conçus dans cet objectif, avec plus ou moins de réussite.

« The dream of having computers able to sustain free conversations with humans has began to become a reality in the last few years, when, thanks to significant advances in the fields of speech and language processing and to the advent of fast processing machinery, the first working prototypes of spoken dialogue systems have been built » [GIA 97].

Toutes les recherches allant dans le sens de l'amélioration de la communication homme-machine 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, elles 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 » [NIC 99].

L'homme, par nature, a la faculté de computer¹, parfois sans même savoir quels sont les mécanismes en jeu dans ses raisonnements. Jean-Louis Le Moigne explique « computer, ce n'est pas seulement calculer arithmétiquement des nombres, c'est, très généralement, manipuler et traiter des symboles » [LEM 86].

Si l'on admet que « l'intelligence est une computation » comme le dit H.A. Simon² dans la nouvelle édition de *The Sciences of the Artificial*, la question se pose

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

2. Prix Nobel de sciences économiques en 1978, après avoir obtenu en 1975, avec A. Newell, le Prix Turing (équivalent du Prix Nobel pour les sciences de la computation).

alors de savoir si une « intelligence artificielle » est envisageable. La formule « intelligence artificielle » retenue pour désigner une discipline scientifique nouvelle dont l'objet serait la représentation, voire l'explication et la reproduction des comportements intelligents (adaptatifs et téléologiques³) d'un système autonome, a été sélectionnée, parmi d'autres expressions possibles, par un jeune professeur de Dartmouth College, J. McCarthy, en 1956. L'intelligence artificielle (IA) débute sans doute en 1847, lorsqu'un logicien et mathématicien anglais du nom de George Boole énonce que « la logique doit être rattachée aux mathématiques et non à la philosophie ». Avec son ouvrage *Une exploration des lois de la pensée, sur lesquelles sont fondées les théories mathématiques de la logique et des probabilités*, il prétend que la pensée humaine est soumise à des lois et que la logique formelle peut décrire ces lois. Ce n'est que quatre-vingt-dix ans plus tard, en 1937, que Claude E. Shannon (à qui l'on doit le « bit » : BInary digiT), mettra en relation, dans sa thèse⁴, les premières formulations de la logique booléenne (et son aspiration à représenter les lois de la pensée humaine) et la représentation matérielle de cette pensée par des circuits électriques.

La première grande avancée dans ce domaine des « automates qui comptent » devait alors être effectuée dans le contexte de la seconde guerre mondiale : Alan Turing avait réussi à mettre au point une machine capable de déchiffrer le code Enigma, qui faisait à l'époque la puissance des armées allemandes. Colossus (c'est le nom donné à cette machine), conçue sous la direction de M.H.A. Newman et T.H. Flowers, à Bletchley Park (Grande-Bretagne) aurait ainsi permis de déchiffrer de 75 000 à 80 000 messages ennemis interceptés.

Claude Shannon et Alan Turing s'intéressaient, chacun de leur côté, à des machines capables de résoudre des problèmes typiquement non numériques. Ils étaient en correspondance depuis 1943, et se rencontrèrent en 1950, pour s'apercevoir qu'ils travaillaient, isolément, sur le même type de problème : la conception d'une machine qui joue aux échecs⁵.

Mais la question la plus intéressante reste celle du langage. « L'homme le plus inculte et le plus hébété comprend naturellement ce que vous lui dites. Comment fonctionne ce système de traitement de l'information qui se montre capable,

3. Fondé sur l'idée de finalité, la téléologie est une doctrine qui privilégie l'explication finaliste.

4. Intitulée *A symbolic analysis of relays and switching circuits* (De l'analyse symbolique des relais et des circuits de communication).

5. Selon Jean-Louis Le Moigne, l'article que publia C. Shannon en février 1950 dans le populaire *Scientific American*, sous le titre *Une machine qui joue aux échecs*, est aujourd'hui considéré comme la référence initiale de nombreux travaux ultérieurs sur les programmes de jeux d'échecs qui ont jalonné, pendant près de vingt ans, les premiers pas de l'intelligence artificielle et des sciences de la cognition.

manipulant intelligemment des symboles, de produire du sens ? » [LEM 86]. La génération du sens demeure aujourd'hui encore un domaine de recherche qui suscite un intérêt très particulier puisqu'il touche à la fameuse notion de « conscience artificielle ». Cardon, dans son ouvrage *Conscience artificielle et systèmes adaptatifs*, décrit l'enjeu de la sorte : « Le problème du sens est le suivant : quels sont les caractères d'un système artificiel qui générerait du sens à propos des choses de son environnement ? » [CAR 99]. Cette obsession de la compréhension de l'extraordinaire problème [MOR 84], la relation de l'esprit et du cerveau, la représentation intelligible de l'activité cérébrale allait guider les chercheurs vers les développements des sciences de la cognition et de l'IA, en posant cette grande question : « qu'est-ce que l'intelligence ? », puis par extension, « les machines peuvent-elles penser ? ». Et pour évaluer cette hypothèse, les scientifiques ne cherchent pas à tester la puissance calculatoire des machines qu'ils mettent au point, mais bien plus encore, leurs facultés à entretenir une conversation avec un interlocuteur humain (voir à ce propos le fameux « test de Turing »).

Le problème le plus délicat, que d'autres avant nous ont déjà mis en lumière, c'est sans doute que l'homme veut tenter de transmettre quelque chose (une intelligence ?) à la machine, alors même qu'il ne maîtrise pas complètement la nature et les phénomènes qui régissent son fonctionnement naturel. Dans son *Histoire universelle des chiffres*, Georges Ifrah livre sa réflexion sur ce sujet, avec un paragraphe qu'il intitule « Lorsque la métaphore fut identifiée à la réalité », et dont voici un large extrait :

« Dans leur enthousiasme, ces savants crurent d'abord que toute opération de la pensée humaine était exclusivement de nature calculatoire (c'est-à-dire algorithmique), et donc que tout processus intellectuel était exécutable sur une machine de type ordinateur.

Et comme ils désignèrent sous le nom de neurones et de synapses des modules qui n'eurent en réalité qu'une très lointaine parenté avec les composants réels, beaucoup plus complexes, d'un encéphale vivant, ils établirent inévitablement, selon un anthropomorphisme simpliste, et forcément réducteur, un parallèle, très étroit entre les circuits d'un ordinateur électronique avec les cellules nerveuses et les composants neurobiologiques du cerveau vivant.

Et c'est ainsi que l'organe central d'un simple ordinateur électronique fut désormais considéré comme l'équivalent du cerveau humain, le public assimilant dès lors les activités d'un ordinateur à celles, éminemment supérieures, de l'esprit humain. D'où, par identification abusive de la métaphore avec la réalité, et par une sorte de convergence avec le mythe du « cerveau électronique », le développement de l'idée d'une machine prétendument douée de pensée inductrice, et créatrice, munie de la faculté de prendre toutes sortes d'initiatives, et à laquelle l'homme pourrait confier tous ses problèmes sans exception pour obtenir d'elle toutes les solutions voulues quasi instantanément » [IFR 94].

Notre propos est donc plus orienté vers l'étude de procédés et technologies permettant à un ordinateur d'imiter la computation humaine de symboles, plutôt que vers celle consistant à inculquer un raisonnement humain (voire une conscience) aux ordinateurs. Nous n'en avons d'ailleurs ni les moyens ni la prétention. Ces deux démarches peuvent sembler relativement proches, mais en fait, la nôtre n'implique pas forcément que la technique utilisée par la machine soit exactement celle employée par l'homme. « Nous devons dès à présent sortir de l'idée que le dialogue homme-machine doit être et peut être identique au dialogue homme-homme car il s'agit là d'une véritable chimère » [LEM 00]. Ainsi, même si la manière de computer des symboles afin d'accomplir une tâche n'est pas la même de part et d'autre, mais que le dialogue reste cohérent pour chacun des interlocuteurs et qu'il mène à une solution possible dans le processus de recherche d'informations, nous aurions atteint notre objectif.

7.2. Dialogue homme-machine sur le web

Le principe que nous suivons consiste à essayer de faire accomplir à la machine, une partie de ce que nous pensons que le cerveau fait lorsqu'il traite de l'information pour lui donner du sens. Selon nous, du point de vue dialogique, tout se passe comme si simplifier, c'était abstraire. C'est-à-dire « passer des réalités sensibles aux concepts, pour les rendre intelligibles » (*Dictionnaire encyclopédique de la langue française*). Pour tenter de comprendre un énoncé, nous ne travaillons pas sur la suite d'éléments qu'il contient, mais plutôt sur les concepts auxquels ils sont généralement associés, dans ce contexte.

Dans le cadre de nos travaux en RI, grâce au TALN 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 concepts⁶ au sein du discours. En effet, nous pensons comme Gérard [GER 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 ».

Il est donc entendu que la compréhension et l'emploi de la LN par la machine seraient très utiles à l'homme, pour la plupart des tâches complexes à (faire) exécuter. L'utilisateur, s'exprimant naturellement, aurait moins d'efforts à fournir pour communiquer avec la machine.

6. 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 (exemple : le concept du temps).

« En laissant à la machine le travail d'interprétation de leurs énoncés, et, à un niveau supérieur, de reconnaissance de leurs intentions, la charge cognitive des utilisateurs est diminuée d'autant. C'est en cela que réside l'intérêt principal du langage naturel en communication : dialoguer en langue naturelle demande peu d'effort conscient aux utilisateurs humains, la charge cognitive étant en quelque sorte déplacée de l'utilisateur vers la machine » [LEM 00].

Pour progresser vers notre objectif de facilitation d'accès aux ressources du réseau internet, grâce à des techniques langagières, il est nécessaire de recueillir des corpus *via* le World Wide Web où le DHM est réel (et crédible), avec un temps de réponse aussi rapide que possible. 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 [ROU 99b]. 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 [NIG 92] ; BOU 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.

La notion d'hyperdialogue que nous venons de présenter est censée améliorer, dans une certaine mesure, la RI d'un utilisateur, en lui permettant d'exprimer ses besoins grâce à un dialogue en LN française, avec la machine. Les domaines d'application sont multiples (serveur météo, bourse, réservation de places de spectacles, renseignements touristiques...). Pour la mise en œuvre de notre système

d'hyperdialogue, nous avons choisi le domaine de la RI documentaire. La partie suivante présente donc le système Halpin-Documentaire accessible sur l'internet.

7.3. Le système Halpin-Documentaire

Le système que nous présentons ci-dessous est l'une des premières mise en œuvre du concept d'hyperdialogue sur le web. Il se nomme Halpin-Documentaire (Halpin signifiant Hyperdialogue avec un Agent en Langage Proche de l'Interaction Naturelle). Nous détaillerons plus particulièrement ici l'architecture et les composantes orales de dialogues développées.

7.3.1. Présentation du système

Le système Halpin-Documentaire vise à permettre à un utilisateur du World Wide Web de dialoguer avec une machine distante, en LN écrite ou orale, afin d'effectuer une RI 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 les suivantes :

- 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, 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). Il 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 d'autre part celui 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 exclusif : 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).

7.3.2. *Modèle de dialogue*

Le modèle de DHM coopératif et finalisé que nous voulons décrire s'inscrit dans une tâche de RI, au cours de processus interactifs : « *As information seeking and retrieval are interactive processes, we believe that providing a flexible and cooperative human-machine dialogue is a complementary way to improve information retrieval systems* » [BAT 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 elles devra être fournie pour amorcer la RI 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 (considéré 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 multiple 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, il est possible de considérer un dialogue entre A, qui demande de l'information, et B, qui en fournit, de la manière suivante :

$$\text{Dialogue (A,B)} \rightarrow \text{requête (A,B)} + \text{promesse (B, A)} + \text{information (B,A)} \\ + \text{test-satisfaction (A,B)}$$

ou bien :

$$\text{Dialogue (A,B)} \rightarrow \text{offre (B,A)} + \text{acceptation (A, B)} + \text{information (B,A)} \\ + \text{test-satisfaction (A,B)}$$

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 [STE 95, 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* » [HAR 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 type d'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 le cas échéant.

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 » [SMA 98]. Ainsi, une plus grande pertinence de répliques devrait inciter l'usager 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 LN 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 cent réponses, alors que la seconde en donne une 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'usager. Mais bien entendu, plus la conversation paraîtra fluide et naturelle pour les

utilisateurs, plus grande sera la complexité du système de TALN. La gestion du DHM prend alors toute son importance.

7.3.2.1. Gestion 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 [GUY 84]. On note dans [BIL 92] :

« 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 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.

7.3.2.1.1. 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 stables⁷, mais les rôles peuvent être modifiés au cours du dialogue, suivant le contexte et les nouvelles situations qui interviennent.

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

7. 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.

techniquement possible d'écouter et de parler simultanément, alors que c'est impossible sur un forum de discussion 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 communiquer⁸ » [WAT 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 pour 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.

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).

8. 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.

7.3.2.1.2. 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 cherchiez⁹ ? »), 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.

9. Extraits du corpus de dialogue homme-homme « Archives départementales de l'Isère ».

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 = oui salaire ¹⁰ très bien

Tableau 7.1. Exemples d'adaptation de la réponse en fonction des critères de recherche

Le tableau 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 document ». Le but et le type d'utilisateur sont également pris en compte dans les exemples du tableau ci-dessus, mais nous verrons cela plus en détail au paragraphe 7.3.2.1.5.

7.3.2.1.3. Prédire les interventions de l'utilisateur

Etant donné 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 » [GUY 95].

La planification est donc vue, de manière classique, 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.

10. 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 a l'air ».

Dans sa thèse, Lehuen identifie quatre approches distinctes pour une modélisation du DHM : la planification, la structuration, la négociation et l'interaction [LEH 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 » [ALL 80].

Depuis les années 1980, plusieurs modélisations, notamment [ALL 87 ; CAR 90], 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 [LIT 85].

En se basant sur les travaux en compréhension de récits¹¹, 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' » [GUY 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 [WAT 76], 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 de RI 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...

11. Voir les publications de Wilenski à ce sujet.

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 ?

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 RI.

Etant donné 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 » [VIL 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 [ROU 98] et sur des travaux du domaine concernant la même tâche [LEH 97], nous avons recensé un certain nombre d'événements susceptibles d'intervenir au cours du dialogue. Dans certains cas (comme en ouverture de

dialogue), les actions possibles sont limitées, tandis que dans d'autres cas (comme en recherche libre), 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.

A chaque événement, correspond une « fiche » permettant de réagir à l'énoncé formulé par l'utilisateur. Prenons l'exemple de l'événement « 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 1 659 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).

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. 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.

7.3.2.1.4. Interpréter les interventions de l'utilisateur

L'aspect technique, même s'il est très important, ne doit pas nous faire oublier l'aspect cognitif du DHM. Autrement dit, il ne *suffit* pas de reconnaître les mots prononcés, il faut aussi les comprendre, et en déduire l'intention de l'utilisateur. Nous rejoignons ici Austin [AUS 62] et Searle [SEA 69] dans la théorie des actes de langage, qui se fonde sur le concept que « dire c'est faire ». Piaget pense que « le

langage se construit par l'action¹² » [PIA 64]. L'école anglo-saxonne en philosophie du langage, estime, quant à elle, que « le langage construit l'action ». Dans cette théorie, parler c'est agir, produire des actes ; communiquer c'est agir sur l'interlocuteur. Dans la phrase « Pouvez-vous fermer la porte ? », le locuteur ne souhaite pas une réponse de bas niveau comme « oui » ou « non », mais désire réellement faire faire un acte à son interlocuteur. Austin isole trois dimensions de l'énonciation :

– l'acte locutoire : c'est l'action effective de dire quelque chose (articulation et combinaison des sons) ;

– l'acte illocutoire : c'est l'acte réalisé par le fait même de parler, et qui modifie ou précise la situation respective des interlocuteurs (accuser, affirmer, autoriser, avertir, conseiller, critiquer, interroger, menacer, ordonner, promettre...) ;

– l'acte perlocutoire : c'est l'ensemble des conséquences qui découlent indirectement de l'acte de parole. Sa fonction n'est pas contenue directement dans l'énoncé, elle dépend entièrement de la situation d'énonciation. Il est caractérisé en termes d'effets réalisés par l'énonciation chez l'interlocuteur (par exemple : effrayer quelqu'un en le menaçant verbalement).

L'acte perlocutoire est sans doute le plus délicat et le plus complexe à appréhender. Comprendre un énoncé, c'est donc, la plupart du temps, comprendre la situation qui l'accompagne. Citons Martin [MAR 83] par exemple, qui explique que la simple phrase « je reviendrai » prend tout son sens en contexte : c'est une *promesse* à la fin d'une visite amicale, un *compliment* de la part d'un client de restaurant, un *avertissement* à un automobiliste mal garé de la part d'un policier, une *menace* du propriétaire à son locataire n'ayant pas payé son loyer, une *prédiction* consolatrice d'un soldat qui part au front...

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 [ROU 99b]. 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. 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

12. Pour lui, l'enfant construit son langage comme résultat de l'assimilation des actions sur le monde. Appeler telle personne « maman » n'a de sens que dans la mesure où celle-ci accourt au cri de « maman ».

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 », « ciao », ou « bonsoir » seront assimilés à des formes de clôture de dialogue, si cette règle est vérifiée.

7.3.2.1.5. 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 [CAE 92] :

- *directive* pour les phases d'ouverture, d'introduction et de clôture (c'est-à-dire à l'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.

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

7.3.3. 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, environ 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 7.1 ci-après présente l'architecture générale du système Halpin-Documentaire.

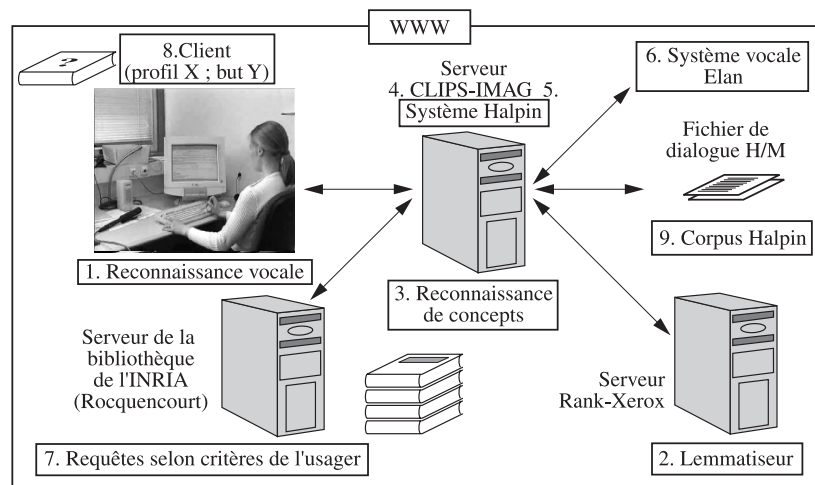


Figure 7.1. Architecture du système Halpin-Documentaire

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 [GAU 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 [ROU 98a, b]. Il fournit une structure de traits à partir d'une analyse conceptuelle de la source lemmatisée ;

4- le *contrôleur de dialogue* a été également développé dans le cadre d'Halpin [ROU 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.

Les éléments importants du système final sont les suivants :

– l'accès à une base de données spécialisée dans le domaine de l'informatique et de l'automatique (INRIA). On utilise le serveur de l'INRIA, situé à Rocquencourt, près de Paris. Cette base de données contient plus de 83 000 références, et demeure en constante évolution ;

– une synthèse vocale de bonne qualité que l'on a intégrée dans une applet Java, et que l'utilisateur peut arrêter ou réécouter à volonté. Le système ne nécessite pas de logiciels supplémentaires (plug-in) pour jouer ces fichiers sonores ;

– 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.

7.3.4. Exemples de dialogues avec Halpin-Documentaire

Voici à présent quelques exemples de dialogues réels entre des internautes et le système Halpin-Documentaire.

(...)
 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.
 (...)

Figure 7.2. Exemple n° 1 de dialogue avec Halpin-Documentaire

Dans le premier exemple (figure 7.2), l'utilisateur change subitement de thème de recherche, en passant de « virus » à « moniteur ». Le système, ne disposant pas d'informations contextuelles, sollicite l'utilisateur pour déterminer si ce nouvel élément de recherche doit être rattaché au slot « auteur » ou, le cas échéant, au slot « thème ». En M10 (M mis pour Machine, H mis pour Homme), le même processus de demande de confirmation grâce au dialogue est utilisé, pour s'assurer que l'élément donné en H9 (feu), est bien relatif au thème recherché.

14:16:23: M0 = Que souhaitez-vous faire ? 1 : Retrouver un document que vous connaissez, 2 : Rechercher un ensemble de documents, ou 3 : Découvrir ce site ?
 14:16:23: H0 = 2
 14:16:36: M1 = Rechercher un ensemble de documents. D'accord. Souhaitez-vous que je vous aide à formuler votre requête ?
 14:16:53: H1 = oui
 14:16:58: M2 = Pouvez-vous me donner le thème recherché, entre guillemets s'il vous plaît ?
 14:17:16: H2 = » linguistique »
 14:17:36: M3 = J'ai trouvé 340 documents ayant comme thème linguistique. Voulez-vous affiner votre requête ?
 14:17:50: H3 = oui

14:17:59: M4 = Afin d'affiner, vous devez choisir parmi les critères : Auteur, Titre, Thème, Année, ou Type de document.
 14:18:31: H4 = année
 14:18:41: M5 = Vous pouvez soit me donner une année, soit effacer l'année en cours.
 14:18:52: H5 = 97
 14:19:08: M6 = J'ai trouvé 10 documents de 1997 ayant comme thème linguistique. Consultez-les sur la page de droite. Est-ce que cela vous convient ?
 14:19:23: H6 = non
 14:19:28: M7 = Souhaitez-vous plutôt affiner, modifier ou annuler votre requête ?
 14:19:43: H7 = modifications-la
 14:19:53: M8 = Vous souhaitez modifier votre requête. Selon quel critère je vous prie ?
 14:20:04: H8 = la date
 14:20:13: M9 = 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 ?
 14:20:24: H9 = 97 ou 98 ou 99
 14:20:47: M10 = J'ai trouvé 10 documents de 1997 ayant comme thème linguistique. Consultez-les sur la page de droite. Est-ce que cela vous convient ?
 (...)

Figure 7.3. Exemple n° 2 de dialogue avec Halpin-Documentaire

Dans le deuxième exemple (figure 7.3) connaissant le but de l'utilisateur (donné en H0) ainsi que son profil (« novice », déduit de la réponse donnée en H1), la machine suggère d'affiner la requête, vu le nombre trop important de réponses obtenues en M3. Elle aide l'utilisateur en M4, en lui rappelant les critères de sélection utilisables dans cette base de données. En M6, après avoir fait l'hypothèse (afin de ne pas alourdir le dialogue inutilement) que 97 était mis pour 1997, le système entre dans un test de satisfaction, et les échanges vont se poursuivre jusqu'à ce que l'utilisateur soit satisfait des documents sélectionnés.

17:49:53: M0 = Que souhaitez-vous faire ? 1 : Retrouver un document que vous connaissez, 2 : Rechercher un ensemble de documents, ou 3 : Découvrir ce site ?
 17:49:53: H0 = 3
 17:50:12: 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.

17:56:01: H1 = 2
 17:56:13: M2 = Rechercher un ensemble de documents. D'accord. Souhaitez-vous que je vous aide à formuler votre requête ?
 17:56:36: H2 = oui
 17:56:42: M3 = Pouvez-vous me donner le thème recherché, entre guillemets s'il vous plaît ?
 17:57:11: H3 = « sémantique »
 17:57:28: M4 = J'ai trouvé 706 documents ayant comme thème sémantique. Voulez-vous affiner votre requête ?
 17:57:39: H4 = oui
 17:57:48: M5 = Afin d'affiner, vous devez choisir parmi les critères : Auteur, Titre, Thème, Année, ou Type de document.
 17:58:22: H5 = thème « ambiguïté »
 17:58:37: M6 = J'ai trouvé 3 documents ayant comme thème sémantique ambiguïté. Consultez-les sur la page de droite. Est-ce que cela vous convient ?
 18:00:08: H6 = tout à fait
 18:00:32: M7 = Très bien, je suis heureux d'avoir pu vous aider dans votre recherche.
 18:01:45: H7 = Merci Au revoir
 18:01:59: M8 = Avant de partir, merci de bien vouloir remplir ce questionnaire s'il vous plaît !

Figure 7.4. Exemple n° 3 de dialogue avec Halpin-Documentaire

Dans ce troisième exemple (figure 7.4), en H5, l'utilisateur donne une information double. Tout d'abord il indique qu'il veut travailler sur le critère « thème », et il précise la valeur de ce thème (le mot « ambiguïté »). Le système comprend parfaitement cette information, et ajoute ce dernier mot-clé au précédent. On affine donc la requête, en passant de 706 réponses à 3 seulement.

7.4. L'hyperdialogue de demain

Nous avons donc 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 [MEI 98] et dans celui du DHM [LEM 00].

Nous avons vu que la question du sens est primordiale, tant dans le domaine de la RI que dans celui des IHM. En RI, bien entendu, si l'on souhaite retrouver et présenter à l'utilisateur une bonne information, il faut tout d'abord comprendre le sens de sa requête. Et en IHM, pour qu'il y ait interaction efficace entre les agents, il est nécessaire qu'un langage, plus ou moins élaboré, soit mis en œuvre. Icône, métaphore, dialogue en LN, son tridimensionnel ? Tout dépend du type d'utilisateur, de la tâche à effectuer, et du domaine visé par l'application. On comprend bien alors que ces deux domaines (RI et IHM) sont étroitement liés. Le plus puissant des systèmes de RI demeure inefficace si ses modes d'entrée et de sortie sont inadaptés. De même, une excellente interface est totalement inutile en soi, si elle n'est pas couplée à un noyau fonctionnel adéquat.

Pour notre part, nous nous sommes intéressé à la recherche d'information documentaire sur le réseau internet. Pour aller plus loin que les traditionnels moteurs de recherche à base de mots-clés et à interaction limitée, nous avons proposé un système de DHM sur le web reposant sur le concept d'hyperdialogue. Nous avons montré comment il était possible, partant d'un système volontairement limité, de l'enrichir de nouvelles connaissances, au fur et à mesure de son utilisation par les internautes, et cela dans un cycle de développement incrémental.

Le corpus de véritables dialogues ainsi obtenu a permis d'étudier plusieurs phénomènes intéressants et nous a apporté des éléments pour la construction des différents modèles utilisés dans ce système : celui de la tâche, de l'utilisateur, du dialogue, et de la langue. L'architecture modulaire mise en place au cours de ce projet a permis de tester différentes versions du système. Outre l'utilisation de briques logicielles externes, comme l'analyseur morphosyntaxique de Xérox, ou la base de données de l'INRIA, cela nous a permis d'intégrer et d'évaluer certains éléments susceptibles d'améliorer l'ensemble du processus d'interaction :

- une vue en œil de poisson, dite Fish-Eye-View, a été intégrée dans le système Halpin-Documentaire [ROU 99c]. 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 colline. On peut également déplacer la souris tout en maintenant le bouton gauche appuyé, et la surface ainsi focalisée se déplace de manière fluide ;

- une ressource linguistique supplémentaire a été apportée au système, en couplant le module de RI à un thesaurus [DUQ 00]. La maquette réalisée utilise le moteur de recherche de Lycos (www.lycos.com) pour retrouver les informations et accède au thesaurus de la Base Terminologique de Québec [BTQ], pour l'aide à la recherche thématique. L'utilisateur peut préciser le type de documents recherché :

texte, image, son, vidéo, et les informations peuvent être retrouvées dans une langue voulue : français, anglais, espagnol, ou toutes les langues disponibles ;

– nous avons également réalisé l'intégration d'un agent virtuel animé dans notre système de dialogue sur le web [ROU 00b]. Selon les premières études, l'utilisateur semble être davantage captivé par un DHM dans lequel il perçoit son interlocuteur. Les « temps morts » comme lors des appels à des ressources distantes *via* le réseau, où d'ordinaire l'utilisateur s'impatiente rapidement, semblent être mieux acceptés lorsque l'agent animé prend le relais pour expliquer ce que fait le système. L'insertion dans une page web d'agents animés tels que ceux proposés par Microsoft [MSA] est, *a priori*, un moyen simple et puissant de donner un corps et un visage (une vie ?) à un agent artificiel. Les hyperdialogues de demain seront sans nul doute basés sur des représentations d'agents virtuels de ce type.

Les interfaces homme-machine (IHM) utilisent de plus en plus de multimodalité. Le dialogue ne se fait plus seulement avec des canaux de communication traditionnels (clavier/souris en entrée, et écran en sortie), mais avec plusieurs (observation par caméra, geste de désignation, objets à retour d'efforts, etc.).

L'utilisation de la voix est un des moyens privilégiés d'interactions avec un ordinateur, puisqu'il semble être – pour l'humain – plus naturel, et demande *a priori* moins d'effort d'apprentissage. Des systèmes de Dialogue homme-machine (DHM) utilisant une reconnaissance et une synthèse vocale existent déjà, mais très peu sont actuellement disponibles pour le grand public, sur le web notamment.

L'émergence de nouveaux langages à balises, comme VoiceXML [VXM] et NLSML [NLS], proposés par le W3C (*World Wide Web Consortium*) pour l'utilisation de la langue naturelle orale sur le web, témoigne de l'engouement que suscitent ces IHM basées sur une composante langagière. Face à ces outils, les orientations sont multiples, et il convient de spécifier, de proposer et d'évaluer l'intérêt et l'apport réel de chacun d'entre eux, compte tenu des besoins et des différents profils des usagers.

De nombreuses pistes de recherche dans le domaine de la RI, du DHM et de l'hyperdialogue s'ouvrent à nous. L'apport d'un langage tel que le VXML, réside principalement dans la souplesse de traitement du dialogue : il est possible de définir une grammaire de type BNF¹³ (Backus Naur Form) pour le dialogue envisagé. On peut obtenir des données provenant de l'internet grâce à un simple téléphone, soit en utilisant les touches du clavier téléphonique (technologie DTMF : *Dual Tone Multi Frequency*), soit directement en parlant à la machine distante. On peut également stipuler que le dialogue sera de type « Barge-in » (interruptible), c'est-à-dire que l'on autorise les interlocuteurs à parler simultanément ou à s'interrompre

13. JSGF (*Java Speech Grammar Format*) est utilisée pour VXML par exemple [JSG].

mutuellement. Cela donne une puissance langagière indéniable au dialogue. L'utilisateur n'est pas obligé d'attendre la fin d'une synthèse vocale pour s'exprimer, par exemple.

Enfin, à terme, la multimodalité devrait être supportée par les navigateurs permettant d'accéder à l'information disponible sur le web. Les spécifications de langages à composantes multimodales, tels que MPML (*Multimodal Presentation Markup Language*) sont en cours. Ils devraient apporter une plus grande liberté d'expression avec de multiples possibilités d'interaction, en entrée aussi bien qu'en sortie de données. Mieux communiquer avec nos systèmes pour être mieux compris, et donc mieux servis, tel est l'enjeu des travaux dans le domaine de la recherche d'information couplée aux interfaces homme-machine intelligentes.

7.5. Bibliographie

- [ALL 87] ALLEN, J.F., LITMAN, D.J., « A Plan Recognition Model for Subdialogues in Conversations », *Cognitive Science*, 11, 2, 1987.
- [ALL 80] ALLEN J.F., PERRAULT, C.R., « Analyzing Intention in Utterances », *Artificial intelligence*, n° 15, 1980.
- [ANI 98] ANICK P. TIPIRNENI, S., « Interactive document retrieval using faceted terminological feedback », *ACM SIGIR'98 Post-Conference Workshop on Query Input and User Expectations*, Melbourne, Australia, 1998.
- [AUS 62] AUSTIN, J.L., *How to do things with words*, Oxford, Clarendon Press, Hermann, traduit sous le titre *Quand dire c'est faire*, Seuil, 1970.
- [BAT 95] BATEMAN, J.A., HAGEN, E., STEIN, A., « Dialogue modelling for speech generation in multimodal information systems », in P. Dalsgaard *et al.* (dir.), *ESCA Workshop on Spoken Dialogue Systems – Theories and Applications* (pp. 225-228), Aalborg, Danemark, ESCA/Université d'Aalborg, 1995.
- [BEL 99] BELL, L., GUSTAFSON, J., « Interaction with an Animated Agent in a Spoken Dialogue System », *Eurospeech'99*, Budapest, p. 1159-1162, 1999.
- [BIL 92] BILANGE, E., *Dialogue personne-machine : modélisation et réalisation informatique*, Paris, Hermès, 1992.
- [BOL 80] BOLT, R.A., « Put-that-here : voice and gesture at the graphic interface », *Computer Graphics*, 14, 262-270, 1980.
- [BOU 92] BOURGUET, M-L., Conception et réalisation d'une interface de dialogue personne/machine. Thèse, Grenoble, septembre 1992.
- [BTQ] Base Terminologique de Québec : <http://www.olf.gouv.qc.ca>
- [CAE 92] CAELEN, J., Compte-rendu du « workshop » interfaces homme-machine multimodales, GDR-PRC CHM sur le dialogue, Dourdan, 1992.

- [CAR 90] CARBERRY, S., *Plan recognition in natural language*, MIT Press, Cambridge, Mass., 1990.
- [CAR 99] CARDON, A. *Conscience artificielle et systèmes adaptatifs*, Eyrolles, 1999.
- [CHA 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.
- [COU 86] COULON, D., KAYSER, D., « Informatique et langage naturel : Présentation générale des méthodes d'interprétation des textes écrits », *TSI : Technique et Science Informatique*, vol. 5. n° 2, 1986.
- [DUQ 00] DUQUE M.V, L'interaction multimodale pour la recherche d'information, DEA Systèmes d'information, Laboratoire CLIPS-IMAG, 2000.
- [GAU 97] GAUSSIER 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.
- [GER 98] GERARD F., NICOLLE, A., « Bistro, un modèle de dialogue intégrant la manipulation de concepts », *Rencontre des Etudiants Chercheurs en Informatique pour le Traitement Automatique des Langues, RECITAL'98*, Le Mans, septembre 1998.
- [GIA 97] GIACHIN E., MCGLASHAN S., *Spoken Language Dialogue Systems*, Chapitre 3 : « Spoken Language Dialogue Systems in Corpus-based methods in language and speech processing », Editions Young et Bloothoof, 1997, p. 69-117.
- [GUY 95] GUYOMARD M., NERZIC P., SIROUX J., « Plans, métaplans et dialogue », *Actes des journées de la communication parlée et dialogue*, ENST, 1995.
- [GUY 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.
- [HAR 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, mai 1996.
- [IFR 94] IFRAH, G., *Histoire universelle des chiffres*, Robert Laffont, 2 tomes, Paris, 1994.
- [JSG] <http://java.sun.com/products/java-media/speech/forDevelopers/JSGF/>
- [LEM 86] LE MOIGNE, J.L., *Genèse de quelques nouvelles sciences : de l'intelligence artificielle aux sciences de la cognition*, Fayard, 1986.
- [LEH 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.
- [LEM 00] LEMEUNIER, T., L'intentionnalité communicative dans le dialogue homme-machine en langue naturelle, Thèse de doctorat d'informatique, Université du Maine, 2000.

- [LEM 99] LEMEUNIER, T., « La modélisation des attentes dans le système AMI », *Actes du colloque RECITAL'99*, Institut d'Etudes Scientifiques de Cargèse (Corse), p. 439-443, 1999.
- [LIT 85] LITMAN, D.J., *Plan recognition and discourse analysis : an integrated approach for understanding dialogues*, Thèse de l'université de Rochester, 1985.
- [MAR 83] MARTIN, R. *Pour une logique du sens*, Presses universitaires de France, Paris, 1983.
- [MEI 98] MEIYE, J.P., ROUILLARD, J., VAUFREDAZ, D., *Webcompletion – protocole de normes associatives sur internet*, Ecole thématique CNRS – Sémantique, Caen (Asnelles-sur-mer), 1998.
- [MIC] <http://msdn.microsoft.com/msagent>
- [MOR 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, Seuil, 1986.
- [MSA] <http://msdn.microsoft.com/msagent>
- [NIC 99] NICOLLE, A., LUZZATI, D., « La langue dans l'interaction personnes/machines », *Atelier thématique de TALN'99*, Cargèse (Corse), 1999.
- [NIG 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.
- [NIG 92] NIGAY, L., COUTAZ, J., « Interfaces multimodales et architectures : fusion et parallélisme », *IHM'92. Quatrième journées sur l'ingénierie des interfaces homme-machine*, ENST, Paris, décembre 1992.
- [NLS] <http://www.w3.org/TR/nl-spec/>
- [PIA 64] PIAGET, J. *Development and Learning. Piaget Rediscovered*, Ripple R.E. and Roccastle V.N. (dir.), School of Education, Cornell University, Ithaca, New York, 1964.
- [ROU 00] ROUILLARD J., *Hyperdialogue sur internet. Le système HALPIN*, Thèse de doctorat d'informatique, Université Grenoble I, 2000.
- [ROU 00] ROUILLARD, J., « Hyperdialogue avec un agent animé sur le web », *ERGO-IHM'2000*, Biarritz, Octobre 2000.
- [ROU 99] ROUILLARD, J., « Le système HALPIN : Hyperdialogue avec un Agent en Langage Proche de l'Interaction Naturelle », *TALN'99*, Cargèse, Corse, 1999.
- [ROU 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, p. 3-20.
- [ROU 99c] ROUILLARD, J., « Navigation versus dialogue sur le web. Une étude des préférences », *IHM'99*, Montpellier, 1999, p. 53-56.

- [ROU 98] ROUILLARD, J., « Hyperdialogue homme-machine sur le World Wide Web : Le système HALPIN », *Colloque international Ergonomie et Informatique avancée ERGOIA '98*, Biarritz, 1998.
- [ROU 98b] 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.
- [SEA 69] SEARLE, J.R., *Speech Acts*, Presses universitaires de Cambridge, 1969. Traduit sous le titre *Actes de langage*, Hermann, 1972.
- [SMA 98] SMAÏL, M., « Vers des systèmes évolutifs de recherche d'information : un état de l'art », *Revue TSI : Technique et Science Informatique*, vol. 17, n° 10, décembre 1998.
- [STE 95] STEIN, A., MAIER, E., « Structuring collaborative information-seeking dialogues », *Knowledge-Based Systems*, 8 (2-3), 1995, p. 82-93.
- [STE 97] STEIN, A., GULLA, J. A., MÜLLER, A. et THIEL, U., « Conversational interaction for semantic access to multimedia information », in M.T. Maybury (dir.), *Intelligent Multimedia Information Retrieval*, Menlo Park, AAAI/The MIT Press, 1997, p. 399-421.
- [VER 97] VERNIER F. et NIGAY L., « Représentations multiples d'une grande quantité d'information », *IHM'97*, Poitiers, France, septembre 1997.
- [VIL 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.
- [VXM] <http://www.w3.org/TR/2000/NOTE-voicexml-20000505/>
- [WAT 80] WATZLAWICK P., *Le changement. Eléments de communication thérapeutique*, Paris, Seuil, 1980 ; 2° éd., collection Points, 1986.
- [WAT 76] WATZLAWICK, P., *La réalité de la réalité, confusion, désinformation, communication*, traduit par Roskis E., Seuil, 1976, 2° éd. 1978.
- [WEI 66] WEIZENBAUM, J., « Eliza – a computer program for the study of the natural language communication between man and machine », *CACM*, n° 9, 1966.