

## Chapitre 8<sup>1</sup>

# Dialogue homme-machine et recherche d'information

### 8.1. Position du problème

Le dialogue en langage naturel est pour la recherche d'information interactive une alternative intéressante pour accéder à des informations, notamment sur le WEB, de plus en plus diverses et disséminées. Entre des moyens d'accès, limités bien souvent à des moteurs de recherche à base de mots-clefs, et la navigation, il y a place pour des solutions comme le dialogue homme-machine qui permettent de *négoier* une requête et d'être aidé au cours de la recherche. La langue naturelle peut apporter facilité, souplesse et puissance d'expression pour formuler sa demande, et l'interaction fondée sur le dialogue permet de rationaliser la tâche de recherche. C'est pourquoi, ce chapitre sur le dialogue homme-machine (DHM) en langage naturel (LN) prend place dans ce livre.

On abordera tout d'abord les problèmes généraux du dialogue, puis quelques aspects de la recherche d'information pour lesquels une interface de dialogue pourrait être utile et enfin une réalisation dans cette voie, le système HALPIN. Le système ayant été évalué auprès d'utilisateurs, les résultats permettront de dresser un premier bilan et de poser quelques perspectives.

---

<sup>1</sup> Chapitre rédigé par JEAN CAELEN.

## 8.2. Dialogue homme-machine

### 8.2.1. La langue naturelle en interaction homme machine ?

L'usage de la langue naturelle (LN) en communication homme machine est à la fois un avantage et un obstacle par rapport aux autres moyens de communication :

(a) c'est un avantage, dans la mesure où le locuteur n'a pas d'effort particulier d'apprentissage à faire pour acquérir de nouvelles habiletés pour communiquer (il est sensé posséder et dominer sa langue maternelle) contrairement aux autres moyens de communication tels que le clavier, la souris, etc. ; il a donc seulement besoin d'un apprentissage pour acquérir des savoir-faire liés à la tâche proprement dite (comme connaître le langage d'interrogation), mais,

(b) c'est un obstacle, dans la mesure où la machine n'a qu'une compréhension limitée du langage naturel. Cette limitation engendre toute une série d'artefacts dans la communication que l'utilisateur doit compenser ; en particulier, il doit mémoriser toutes les restrictions imposées par le concepteur au langage d'interaction, et il doit être entraîné aux comportements de la machine en réponse à ses énoncés. En particulier la machine n'a pas la capacité d'anticipation ni la compréhension des intentions qui permettent toutes les richesses que l'on observe dans le langage utilisé par les humains. En général, l'utilisateur ne pourra pas parler par sous-entendus, ni introduire facilement de nouveaux mots, ni construire des sens dérivés, etc. toutes choses que l'on fait spontanément en parlant à quelqu'un.

Il faut donc s'interroger sur les aspects de la communication en langue naturelle qu'il faut impérativement conserver, pour la rendre *attrayante* et *compétitive* vis-à-vis des autres modes d'interaction.

(a) Ce qui rend attrayant le langage naturel n'a pas vraiment été étudié en détail en situation opérationnelle (autour d'une tâche). On sait seulement que le langage opératif [OZK 93, 94] doit conserver son pouvoir constructif ainsi qu'un minimum de spontanéité pour rester utile en communication homme machine. On sait aussi que le langage apporte des modalités temporelles (réitération d'actions, renvoi d'une action dans le futur), et des expressions référentielles condensées (dénominations construites en cours de tâche, ellipses, désignations anaphoriques) dont on ne peut vraiment se passer. Ce qui rend la langue naturelle attrayante dépend donc de ses capacités d'expression.

(b) Ce qui la rend compétitive par rapport aux autres modes de communication doit être examiné de manière comparative pour extraire les modes qui seraient les plus adéquats à la tâche. Des études sur l'interaction multimodale apportent quelques éléments à cette question [CAT 95], [ZAN 97]. On sait par exemple que des situations sémantiquement riches ou des tâches à

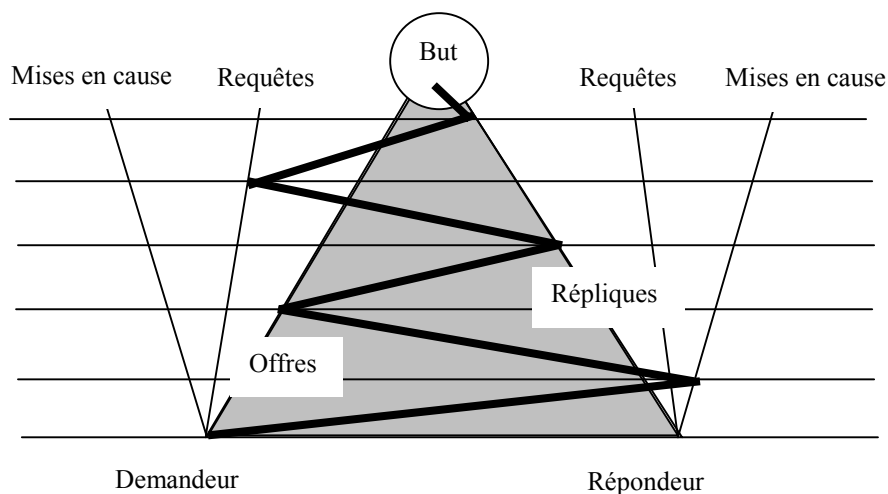
séquencement complexe font davantage appel aux ressources du langage naturel. Les objets non perceptibles, les actions différées, les actions conditionnelles, etc. sont aussi plus facilement exprimables en LN. Mais aussi, et surtout, avec le LN il est possible de former de nouveaux concepts et de construire de nouveaux énoncés.

### 8.2.2. *Le langage, le dialogue et l'action*

« *Le langage se construit par l'action* » : c'est un des résultats principaux de Piaget en examinant l'apprentissage du langage chez les enfants [PIA 64]. Pour lui, l'enfant construit son langage comme résultat de l'assimilation des actions sur le monde. Appeler cette personne « maman » ne fonctionne que si celle-ci accourt au cri de « maman ».

Réciproquement « *Le langage construit l'action* ». C'est la thèse principale de l'école anglo-saxonne en philosophie du langage [AUS 62], [SEA 72]. Dans cette théorie, parler c'est agir, produire des actes ; communiquer, c'est agir sur et *via* l'interlocuteur. Cette théorie permet donc de considérer le langage comme une forme d'action et, par généralisation, de considérer le dialogue comme une séquence d'actions planifiées ayant pour objectif un *but visé sous-tendu par une intention*. Cette conception présume qu'il existe un *équilibre rationnel* entre les connaissances, actions et intentions du locuteur. Le principe de sincérité pose par exemple, que la série d'actions que l'utilisateur est en train de faire, coïncide avec la réalisation de ses intentions, et qu'en vertu des conditions de succès attendu, l'utilisateur a seulement des intentions qui sont possibles à réaliser [SEA 83, 85]. Pour ces philosophes, un acte de langage s'exprime par F(p) où F est la force illocutoire et p le contenu propositionnel. F sous-tend un but illocutoire, par exemple faire-faire quelque chose à quelqu'un (directif). Vanderveken [VAN 90, 91] énonce ensuite une série de conditions qui font qu'un acte sera atteint (conditions de réussite) et satisfait (conditions de succès).

« *Le dialogue est une interaction : il renvoie le langage à l'action et réciproquement* ». Le dialogue est une suite coordonnée d'actions (langagières et non-langagières) devant conduire à un but [VER 92]. Le dialogue est donc un projet conjoint : il avance vers un but commun et tend à réduire les écarts initiaux entre les buts des interlocuteurs. Ceux-ci agissent sous forme de requêtes, répliques, réponses, mises en question ou même remises en cause : ce sont des actes conjoints (fig. 8.1). Au cours de cette interaction ils modifient leurs connaissances, leurs croyances, acquièrent éventuellement de nouvelles connaissances tant sur la situation, que sur leur interlocuteur ou sur la langue.



**Fig. 8.1.** : *Modèle projectif du dialogue [VER 92]. Dans le cas d'un dialogue réussi, le dialogue converge sur le but à atteindre. Ce schéma – qui représente un Echange – montre les axes convergents (Offres et Répliques) et les axes divergents (Mises en cause et Requêtes) du dialogue.*

« *Le sens se négocie au cours du dialogue* », c'est une évidence, notamment le sens général d'un mot s'affine en cours du dialogue, se négocie entre les interlocuteurs. Voici un exemple tiré de [LUZ 89] :

- D : Pouvez-vous me donner l'adresse d'un taxidermiste ?  
 R : Comment ? Qu'est-ce qu'un taxidermiste ?  
 D : C'est un empailleur d'animaux.  
 R : Ah oui ! Non, je ne connais pas de taxidermiste...

En résumé le dialogue est un processus co-construit par les interlocuteurs qui tendent à satisfaire leurs buts au moyen du langage.

### 8.2.3. Modèles de dialogue

On distingue deux grandes classes de modèles de dialogue :

- les modèles statiques,
- les modèles dynamiques.

Les modèles statiques sont habituellement des modèles structurels. Le modèle genevois [ROU 85] est typique de cette classe de modèles, il hiérarchise les actes de

langage en tours de parole puis en unités de niveaux supérieurs, les interventions et les échanges. Les modèles dynamiques quant à eux, se fondent sur des règles en termes de planification [LIT 87, 90], d'états mentaux [COH 78,79], [ALL 80] ou de jeux [MAU 01]. Notons qu'il existe des modèles mixtes par exemple, le modèle de Luzzati [LUZ 95] introduit une dimension dynamique au modèle genevois à l'aide du concept d'axe régissant et d'axe incident du dialogue (l'axe régissant est celui qui gouverne les buts de la tâche et l'axe régissant est celui qui canalise le dialogue).

Avant de décrire les caractéristiques essentielles de ces modèles, il est utile de préciser les fonctions du contrôleur de dialogue dans un système interactif et le cadre pragmatique dans lequel se situe la modélisation du dialogue.

#### **8.2.4. Fonctions du contrôleur de dialogue**

Le contrôleur de dialogue est une sorte de plaque tournante de haut niveau dans les systèmes interactifs. Il est normalement chargé :

- de la construction d'un univers sémiotique partagé et de l'échange des connaissances,
- de l'organisation du dialogue (gestion des tours de parole, des échanges, des interventions),
- du choix des stratégies de dialogue,
- de la réparation des erreurs de communication et du maintien de la communication (relances, phatiques, etc.),
- des aides dans la tâche et dans la conduite des activités.

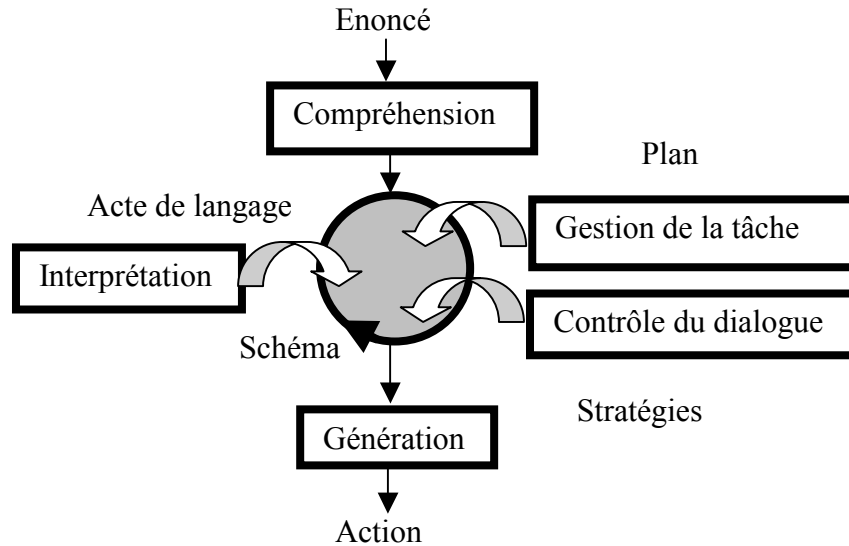
Ce contrôleur de dialogue est rarement lui-même dynamique. Ce qui peut être dynamique est donc seulement l'organisation de la communication et le contrôle de l'interaction, ceci en mettant en jeu diverses stratégies et tactiques de dialogue. Dans un système interactif, le contrôleur de dialogue gère les échanges entre divers modules spécialisés qui sont généralement :

- le module de reconnaissance et de compréhension de la parole (ainsi que d'autres modalités sensorielles s'il y en a),
- le module d'interprétation<sup>2</sup> des actes dialogiques dans le contexte de la tâche en cours,
- le module de gestion des buts (tant les buts actionnels que les buts conversationnels) liés à la tâche proprement dite,

---

<sup>2</sup> La différence entre interprétation et compréhension est de nature pragmatique vs. sémantique : la compréhension s'attache à la sémantique des énoncés, l'interprétation s'attache aux actes de langage contenus dans les énoncés (force illocutoire, référents, etc.), pour simplifier, l'un produit le sens littéral et l'autre le sens contextuel.

- le module de génération des réponses (texte, discours, etc.).



**Fig. 8.2. :** Décomposition fonctionnelle d'un système de dialogue. Un énoncé est analysé sous forme d'acte de dialogue et interprété en regard de la situation, essentiellement le plan d'action et les connaissances d'arrière-plan. Cet acte est ensuite projeté dans un schéma par rapport auquel il est défini en "compréhension". Le contrôleur du dialogue détermine la meilleure stratégie d'interaction et réagit sous forme d'action (langagière ou non).

Le contrôleur de dialogue gère également les données communes à tous ces modules (et évidemment les siennes propres). Cette structure commune est parfois représentée à l'aide d'une CMR (Common Meaning Representation) qui s'enrichit à la manière d'une structure de *tableau noir*. Nous donnerons à cette CMR le nom français de *schéma*.

Définitions :

**Schéma** = structure de données commune représentant le but et sous-buts de la tâche ainsi que les connaissances mises en jeu pour atteindre le but,

**Plan** = base de scripts sélectionnés (ou appris en cours de dialogue) en fonction des connaissances sur l'utilisateur et des connaissances pragmatiques (mondes d'arrière-plan, situation, etc.), munie d'opérations d'enchaînement, permettant de planifier les actions de la machine pour atteindre le but [BRA 87],

**Stratégie** = ensemble de règles comportementales qui régulent le dialogue,

**Action** = réponse de la machine (éventuellement multimodale) en terme de changement d'état dans la situation et dans les connaissances.

### 8.2.5. Un cadre pragmatique

Les éléments sur lesquels s'appuie le contrôleur de dialogue et que nous pouvons maintenant formuler plus précisément, sont essentiellement des connaissances pragmatiques [ARM 94]. Formellement, et réduite à ce propos, la pragmatique se réduit à :

Pragmatique = {modes, interlocuteurs, monde, contexte opératif, actes}

avec :

modes =  $\mu$  = {parole, geste, etc.}

interlocuteurs = {usager = U, machine = M}

monde =  $m$  = {mondes d'arrière-plan  $m_i$ , situation  $\xi(t)$ }

contexte opératif = {tâche T, activité A}

actes = {communicationnels, transactionnels}

et :

U = {compétence = (langagière, opératoire, connaissances du domaine)

performance = (comportementale, perceptive, habiletés motrices)}

Le dialogue se situe au niveau de la coordination des actes des interlocuteurs. Ces actes modifient la situation, c'est-à-dire font passer à un instant donné la situation  $\xi(t)$  à une nouvelle situation  $\xi(t+1)$ . L'utilisateur doit effectuer une tâche T, c'est-à-dire doit atteindre un certain but B caractérisé par la situation  $\xi_B$ . On suppose que ce but motive et engage l'utilisateur (et maintient le dialogue) tant qu'il n'est pas atteint et satisfait.

#### Les modes

Dans le dialogue homme-machine nous appelons *langue* la composante langagière parlée, doublée, éventuellement, d'une gestuelle déictique ou expressive. Elle peut être formalisée par l'intermédiaire d'une grammaire multimodale. Elle ne sera pas détaillée ici. Un énoncé prend donc le sens général d'*acte* (langagier ou non).

#### Les interlocuteurs

Les locuteurs mis en jeu ici sont typiquement l'utilisateur (U) et la machine (M). Il est utile de classer les usagers en types (profils) pour déclencher les stratégies de dialogue et de sous-dialogues et particulièrement, les aides et les guides de manière appropriée. Par souci de simplification on considère souvent uniquement les trois catégories d'utilisateur suivantes :

U = {expert, occasionnel, novice}

catégories qui résument des types plus complexes décrivant des compétences et performances diverses :

### **Le monde**

Le monde (parfois appelé univers) est caractérisé par les mondes d'arrière-plan c'est-à-dire par l'ensemble des connaissances liées au langage, à la tâche, etc. et par la situation présente et passée (mémorisée dans l'historique du dialogue et l'historique de l'activité).

### **Le contexte opératif**

Le contexte opératif englobe la tâche et l'activité de l'utilisateur déployée à propos de la tâche [BRA 87]. La tâche est de nature prescriptive (ce qu'il faut faire) alors que l'activité est de nature proscriptive (ce qu'il se fait ou ce qui a été fait).

### **Les actes de langage**

Les actes se divisent en deux groupes : les actes *communicationnels* en vue d'informer l'interlocuteur ou la machine — ils ne changent pas directement la situation mais simplement les états mentaux et les connaissances des interlocuteurs — et les actes *transactionnels* à proprement parler — qui changent l'état de la situation (appelés aussi performatifs). Voici quelques exemples d'actes de langage :

- informer, asserter : exprimer comment sont les choses
- répéter, épeler : exprimer comment sont les énoncés
- illustrer, exposer, décrire : exprimer comment sont les objets de la tâche
- confirmer, contester, rectifier, réparer : exprimer comment sont les connaissances
- résumer, expliquer, justifier : exprimer comment sont les buts
- questionner, demander, remercier : exprimer comment sont les échanges
- suggérer, ordonner, commander : exprimer comment sont les actions
- introduire, conclure : exprimer comment est le dialogue

### **Les connaissances**

L'ensemble des connaissances mises en jeu au cours d'un dialogue sont nombreuses : il y a bien sûr les connaissances proprement dites sur la langue mais également sur la tâche, la situation courante, les conventions sociales, ainsi que toutes les connaissances encyclopédiques que partagent (ou non) les interlocuteurs. Ces connaissances peuvent être représentées par des *croyances* à la fois sur le monde représenté à travers la situation, sur les mondes d'arrière-plan et sur eux-mêmes.

Une fois ces éléments généraux précisés, nous pouvons maintenant examiner quelques classes de modèles de dialogue.



### 8.2.6. Le modèle structurel (ou hiérarchique)

Le modèle "genevois" [ROU 79, 85], [MOE, 89, 94] est à la source de la plupart des modèles structurels. Il est souvent utilisé en DHM car il est d'une implémentation claire et commode. De ces emprunts dans le monde informatique, sont nés diverses variantes et améliorations. On peut citer par exemple [LUZ 89], [BIL 92], [VIL 92], [LEH 97] en France.

Ce modèle est de nature hiérarchique et se développe sur plusieurs axes (l'axe régissant et l'axe incident chez Luzzati, directeur et subordonné chez Bilange). Il se caractérise essentiellement par l'existence d'une structure à l'intérieur de laquelle le dialogue peut se tisser. Cette structure peut se décrire à l'aide d'une grammaire hors-contexte. Par exemple Bilange propose la grammaire suivante pour un dialogue finalisé (une application de réservation aérienne) :

$$\begin{aligned} E &\rightarrow I.\{I\}^* \\ I &\rightarrow CD \mid CS.CD \mid CD.CS \\ CD &\rightarrow \{A \mid I\}D \\ CS &\rightarrow \{E \mid I \mid A\}S \end{aligned}$$

avec

Catégories discursives : E = échange, I = intervention, A = acte de langage

Fonctions : S = subordonné, D = directeur

Constituants : CS constituant subordonné, CD constituant directeurs

Une expression telle que  $CD \rightarrow \{A \mid I\}D$  se lit « un constituant directeur est formé d'une suite d'actes directeurs et/ou d'une intervention directrice.

A ces règles structurelles qui régissent l'enchaînement possible des constituants du dialogue s'ajoutent des fonctions que le modèle prévoit (fonctions illocutoires essentiellement). Par exemple les fonctions d'initiative, de réactive et d'évaluative, parfois de conclusive (pour la clôture).

Une initiative ouvre un échange (ou réintroduit un échange momentanément mis en attente), ce qui met l'allocutaire en position de réaction. La plupart du temps cet effet attendu l'amène à produire une réactive, à laquelle le locuteur réagit à son tour par une évaluative ou par une clôture, lorsqu'il a obtenu ce qu'il désirait. Notons que la clôture approbative n'est pas toujours marquée explicitement. Quant à l'intervention évaluative, elle peut être positive (elle marque l'avancée vers le but) ou négative (écartement du but) : il s'agit d'actes qui marquent la position (et parfois les raisons de) du locuteur par rapport à son interlocuteur.

Un échange, pris dans sa globalité, peut être lui-même initiatif (échange de politesses par ex.) ou réactif ou évaluatif ou conclusif. Dans un échange complet il peut y avoir tout ou partie de ces divers types d'échanges.

Exemple (cité de [BIL 92 p.106]):

L : Quand voulez-vous partir ?	(1) Intervention initiative
A : Le 13 novembre	(2) Intervention réactive
L : Le 13 novembre...	(3) Intervention évaluative
A : A quelle heure ?	(4) Intervention initiative
A : Non, le 20 novembre !	(5) Intervention évaluative à (3)
A : 10 heures	(6) Intervention réactive à (4)

Dans cet échange complet, (3)+(5) est un échange évaluatif et (4)+(6) un échange réactif. Ces deux échanges sont imbriqués. La clôture est ici implicite, le locuteur ayant obtenu les renseignements qu'il demandait.

Les limites de ce modèle proviennent de la difficulté d'interpréter les fonctions des actes : (6) est une réaction à (4) car (4) est une demande et (6) une réponse cohérente dans le cadre de la tâche sous-tendue par le dialogue : une demande de renseignement d'horaire de train. On ne peut donc interpréter correctement la séquence qu'au regard de la tâche d'une part, mais aussi au regard du but poursuivi par le demandeur (il s'agit pour lui de fournir un billet en consultant une base de données et non de partir lui-même en train).

### 8.2.7. *Le modèle planification*

Dans cette approche de gestion de dialogue, on ne considère plus seulement la structure et la fonction des actes de dialogue mais également les intentions sous-jacentes des interlocuteurs. Les locuteurs dialoguent pour réaliser un but (c'est particulièrement vrai dans les dialogues finalisés), ils se servent pour cela de plans et de schémas [ALL 84], [GUY 85], [GRO 90]. Cette approche se situe donc dans le cadre plus général de la théorie de la planification [NIL 80]. Par analogie aux actions d'un robot pour atteindre un but, les actes de langage sont des actions faites pour modifier les mondes de connaissance des interlocuteurs et le monde de la tâche. Lorsqu'il s'agit d'un dialogue homme-machine, le rôle de la machine est donc de comprendre le plan de l'utilisateur pour l'aider dans sa tâche. Avant de comprendre le plan du locuteur il s'agit bien sûr de le reconnaître puis d'identifier les buts restant à atteindre et correspondant aux intentions de l'utilisateur. L'approche est donc schématiquement la suivante :

- **si** l'on suppose que le locuteur a des buts, qu'il planifie sa tâche et le dialogue en produisant des actes de langage,
- **alors** la machine doit reconnaître le plan à travers les actes de langage de manière à déduire les buts du locuteur.

Cette approche nécessite une modélisation précise des plans et des buts des interlocuteurs. Les buts sont modélisés par des *opérateurs de croyance* sur les

connaissances (par ex. objets de la transaction dans le cas de la demande de renseignements). Les plans sont modélisés comme des suites d'actions mettant en relation un état initial du monde et un état but. Il s'agit alors d'activer les plans pour atteindre cet état but. Les méthodes mises en œuvre relèvent de la planification en Intelligence Artificielle [WIL 83], [LIT 87], [CAR 88], [NER 92]. On utilise souvent des représentations d'actions et de plans fondées sur des *schémas* comme par exemple :

prendre\_train(Passager, Train)  
 définition : appartient(Voiture, Train)  
 corps : embarquer(Passager, Voiture)  
 pré-conditions : possède(Passager, Billet)  
                   sur(Passager, Quai)  
 effets : dans(Passager, Voiture)  
 contraintes : accoster(Train, Quai)

Les limites d'un tel modèle sont évidentes : on ne peut traiter que des dialogues qui sont dirigées par la tâche et dont le but est connu. Les incidences *hors tâche* du dialogue sont quasiment impossibles à traiter (on peut ici répondre à des questions comme "que faut-il faire avant d'embarquer ?", réponses : "être en possession d'un billet et attendre que le train ait accosté au quai", mais on ne peut pas répondre à une question comme "combien de temps faut-il attendre avant d'embarquer?").

### 8.2.8. *Le modèle mental*

Le modèle mental est une extension du modèle de planification. Il considère que les deux modèles de tâche et de dialogue sont distincts d'une part et il se focalise davantage sur la logique dialogique plutôt que sur la logique de la tâche d'autre part. Les initiateurs de ce type de modèle sont Cohen, Perrault et Allen [COH 79], [PER 80]. Plus récemment Sadek [SAD 91, 96] a raffiné ce type de modèle en définissant les principes d'un *agent rationnel dialoguant*. Cet agent doit avoir des capacités de négociation permettant une très grande flexibilité dans l'interaction : il doit suivre le dialogue en autorisant les incidences, reprises, explications, demandes d'aide, etc. tout en ne perdant pas de vue le but final. Il doit donc à la fois avoir des réactions coopératives et prendre des initiatives dans la tâche, "un système de dialogue intelligent doit en fait, être un système intelligent qui dialogue". Sadek rajoute : "ce système doit avoir un comportement rationnel" [SAD 96 p.288]. Cela l'amène à définir un cadre logique complet apte à rendre compte d'*états mentaux*.

En logique dialogique (LD) [CAE 95] on suppose que les interlocuteurs construisent leurs dialogues de manière rationnelle — c'est-à-dire autour d'actions coordonnées — en respectant des conventions (sociales) normalisées. On suppose en

général que le dialogue est à la fois constructif [SUC 90] (il conduit à la construction d'un but à partir des objectifs des interlocuteurs et à l'enrichissement des connaissances mutuelles) et co-interactif (les acteurs coordonnent leurs actions pour aboutir à un certain but). La LD s'appuie sur plusieurs logiques dont la logique épistémique qui traduit des hypothèses sur les *états mentaux* des conversants en terme de *croyances* et d'*engagement*. Ces états mentaux résument l'état de connaissance (croire), le contexte de l'action (faire), les choix (pouvoir, devoir) et les engagements (vouloir). A chaque instant du déroulement du dialogue, le modèle fait l'hypothèse que le dialogue est dirigé par les états mentaux qui sous-tendent les actes (intentions, choix, engagements, etc.) [COH 90].

L'approche générale de cette théorie est fondée sur la rationalité qui ne peut être mise en œuvre en fait, qu'à travers la conscience. On suppose que les conversants sont conscients des buts qu'ils poursuivent (leurs désirs sont rationalisés), maîtrisent les connaissances dont ils disposent (ils savent les évaluer, les faire partager, les remettre en question, etc.) et respectent les conventions sociales [SHE 77], [GOF 81] à travers les rôles qu'ils jouent (et dont ils ont aussi conscience). Dans le cadre du dialogue humain, on peut trouver une telle théorie très limitative : elle ne fait pas, en effet, la place aux inattendus (la planification et l'intention prennent une place dominante) et place les désirs sous le contrôle de la rationalité (ces désirs sont non seulement rationalisés mais aussi conscients). Mais en DHM cette réduction peut être intéressante dans la mesure où, malgré tout, elle assigne une place au concept d'états mentaux et donc à une composante cognitive dans un dialogue pseudo-naturel. Le contrôleur du dialogue a alors pour mission de gérer les actes de l'interlocuteur reflétant ses états mentaux supposés. Cette modélisation ne vise pas à donner un comportement de type humain à la machine mais seulement à lui fournir des éléments logiques de choix et des raisons d'agir. On ne cherche pas à donner des "états mentaux" à la machine mais à modéliser les raisons (à cause de) d'action de l'interlocuteur. Cependant nous désignons — à la fois au sens de *design* et au sens d'interlocuteur.

### 8.2.9. La pertinence (en théorie des jeux)

Actuellement la théorie des jeux de langage [HUL 00] tend à se développer pour modéliser le dialogue. On suppose que chacun des interlocuteurs est engagé dans un jeu dont les tours de parole représentent des *coups*. Ces coups sont plus ou moins pertinents (c'est-à-dire qu'ils atteignent plus ou moins leur but) [SPE 86]. Le dialogue a toujours pour objet un ensemble d'actions en vue de satisfaire un but (B) mais ici ce but B est explicite ou implicite et il est partagé par les interlocuteurs ou non. Dans le cas de *non partage* on peut supposer que les interlocuteurs poursuivent chacun leur propre but — à savoir  $B_E$  et  $B_D$  — jusqu'à ce qu'une explicitation (ou une négociation), à un certain moment du dialogue, leur permette de réajuster leurs

buts respectifs :  $B_E$  et  $B_D$  peuvent tendre vers  $B$  si la négociation réussit ; le dialogue devient alors coopératif, les deux interlocuteurs poursuivant ensemble le même but  $B$ .

La pertinence d'un acte de dialogue énoncé à l'instant  $t$  par  $E$ , est relative au but poursuivi par  $E$  mais aussi relative pour  $D$ , au but que  $D$  poursuit de son côté. On distingue donc deux cas, (a) *coopération* et (b) *concurrence* :

- dans le cas où le but est partagé,  $B_E = B_D = B$ , la pertinence d'un acte  $\alpha(E,D,t)$  de  $E$  à l'adresse de  $D$ , doit amener  $E$  et  $D$  dans une situation de coopération (ou les maintenir dans cette situation s'ils y étaient déjà) et contribuer à les rapprocher du but  $B$ ,

- dans le cas où le but n'est pas partagé,  $B_E \neq B_D$  on doit distinguer une série de scénarios possibles ;

- (a) soit  $E$  et  $D$  engagent une négociation qui peut réussir ou échouer,

- (b) soit  $E$  et  $D$  restent sur leurs positions sans chercher à les négocier ;

à l'issue de cette phase de dialogue on retombe dans une situation de coopération ou on reste dans une situation de concurrence. La stratégie de  $E$  vis-à-vis de  $D$  est alors d'arriver à  $B_E$  ou d'empêcher que  $D$  n'arrive à  $B_D$ . La pertinence des actions de  $E$  peut prendre alors un sens négatif pour  $D$ .

#### La pertinence transactionnelle

La pertinence telle que nous venons de l'envisager, est une pertinence transactionnelle. Elle prend une valeur comprise entre 0 et 1 et peut être définie comme suit :

- *pertinence constructive* :  $P_C\{\alpha(E,D,t)\} = \exp\{-d[B_E, \text{effet}[\alpha(E,D,t)]]\}$

où  $d$  est une distance<sup>3</sup> qui mesure l'écart entre le but recherché par  $E$  et le sous-but projeté par  $E$  manifesté par l'effet de l'acte  $\alpha(E,D,t)$  ; il suffit de considérer les effets potentiels de l'acte  $\alpha$  pour juger de la pertinence et non de son exécution réelle, puisqu'elle est relative au but de l'énonciateur seul.

- *pertinence obstructive* :  $P_{\sim C}\{\alpha(E,D,t)\} = 1 - \exp\{-d[B_D, \text{effet}[\alpha(E,D,t)]]\}$

ici  $E$  s'oppose à la réalisation du but de  $D$ .  $d$  est une distance qui mesure l'écart entre le but recherché par  $D$  et le sous-but projeté par l'effet de l'acte  $\alpha(E,D,t)$  ; si cette distance augmente, l'acte  $\alpha$  aura été (négativement) pertinent puisqu'il éloigne  $D$  de son but.

Ex : supposons  $B_E = (\forall x) (\text{DOCUMENT}(\text{POESIE}(x), \text{AUTEUR}(\text{Baudelaire})))$

$E$  cherche toutes les poésies de Baudelaire dans une base de données

A l'instant  $t=0$  la zone de requêtes est vide.

<sup>3</sup> La distance peut se mesurer par exemple en nombre de tours de parole pour atteindre le but.

- si  $\alpha(E,D,1) = \text{« donne-moi les poésies de Baudelaire »}$ ,  $P_c = \exp(-0) = 1$  car l'effet attendu de la commande conduit au but B (sans présupposer que le destinataire sache répondre à la requête).
- si  $\alpha(E,D,1) = \text{« donne-moi des ouvrages de Baudelaire »}$ ,  $P_c = \exp(-1)$  car il faudra un tour de parole supplémentaire pour demander les poésies parmi les ouvrages.
- si  $\alpha(E,D,1) = \text{« dessine un triangle »}$ ,  $P_c = 0$  car l'effet de l'acte ne conduit pas au but dans un nombre fini de tours de parole.

### La pertinence informationnelle

La pertinence transactionnelle telle que nous venons de l'envisager, s'accompagne toujours d'une pertinence informationnelle dans la mesure où l'acte  $\alpha(E,D,t)$  est énoncé en langage naturel et nécessite une interprétation exigeant un certain nombre de connaissances de la part des interlocuteurs (dans l'exemple ci-dessus il fallait connaître le type d'écrits de Baudelaire). Il est possible de définir cette pertinence informationnelle sur plusieurs axes :

- *pertinence sémantique* : c'est la *pertinence du dire*, c'est-à-dire la pertinence du bon usage des mots (ou des concepts) pour ce qu'ils signifient. On suppose que l'on dispose d'un lexique sémantique du domaine ou de l'application hiérarchisé en catégories ordonnées (thésaurus). On définit pour chaque mot lexical (ou expression pleine) du langage employé, sa place dans l'arbre des catégories sémantiques pré-définies. La pertinence sémantique d'un énoncé se définit alors par :

$$P_S\{\alpha(E,D,t)\} = \text{moy}\{\exp[-\Delta\text{niv}(\text{mot})]\}$$

où l'opérateur *moy* est l'opérateur *moyenne* portant sur tous les mots lexicaux de l'énoncé et où  $\Delta\text{niv}(\text{mot})$  est la différence de niveau dans l'arbre des catégories sémantiques entre le niveau d'emploi de ce mot et le niveau réellement utile dans l'énoncé.

Par exemple :

(a) « lieu de spectacle » est moins pertinent que « cinéma » si le contexte nécessite l'information « où l'on projette des films », car cinéma est à un niveau de spécialisation plus adéquat pour ce contexte,

(b) « je vois un airbus haut dans le ciel » le mot airbus est trop spécialisé si le mot « avion » peut suffire.

La pertinence sémantique permet de qualifier l'usage des synonymes, hyperonymes, hyponymes, antonymes, et des figures de substitution rhétoriques. Plus généralement la pertinence sémantique se rapporte à l'adéquation des mots ou des concepts sur l'axe paradigmatique.

• *pertinence pragmatique* : c'est la bonne adéquation des mots aux choses dans le monde, c'est-à-dire la bonne utilisation des expressions linguistiques pour référencer des objets. La pertinence pragmatique peut se mesurer par :

$$P_p\{\alpha(E,D,t)\} = \min \{N_R/N_E, N_E/N_R\}$$

où  $N_R$  est le nombre d'objets référencés par l'énoncé  $\alpha(E,D,t)$  et  $N_E$  le nombre d'objets visés par l'énonciateur E.

Par exemple :

- (a) « le musée de Paris »,  $N_R = 10$  (les musées),  $N_E = 1$
- (b) « le musée d'Orsay »,  $N_R = 2$  (nom propre et ville),  $N_E = 1$
- (c) « le musée du Louvre »,  $N_R = 1$ ,  $N_E = 1$
- (d) « le musée de Trifouilly »,  $N_R = 0$  (pas de musée ou ville inconnue),  $N_E = 1$
- (e) « tous les musées de Paris »,  $N_R = 10$  (les musées),  $N_E = 10$

• *pertinence épistémique* : c'est l'adéquation des signifiés aux connaissances du destinataire, c'est-à-dire la probabilité de dire ce qui est nécessaire et suffisant à D (et au moment adéquat) pour qu'il comprenne (maxime de quantité de Grice [GRI 75]). La pertinence épistémique se formule par :

$$P_e\{\alpha(E,D,t)\} = \Pr(I_E/I_S) \cdot \Pr(I_N)$$

où  $I_E$  est la quantité d'informations portée par l'énoncé de E,

$I_S$  est l'information suffisante à D (contenue dans l'énoncé) et  $I_N$  est l'information nécessaire à D pour effectuer l'action  $\alpha(E,D,t)$ .

$\Pr$  est une mesure de probabilité qui peut-être estimée en comptant le nombre d'attributs utiles à une information et en les pondérant selon leur importance.

Par exemple :

(a) « Paris, capitale de la France » si D = adulte cultivé  $\Pr(I_E/I_S) = 1/3$  (le seul terme Paris est suffisant, capitale et France sont redondants),  $\Pr(I_N)$  est la probabilité de parler nécessairement de Paris à ce moment-là. Mais si D = enfant en cours de géographie, alors  $\Pr(I_E/I_S) = 3/3$  et  $\Pr(I_N) = 1$ .

(b) « Paris, province de la Belgique » bien que paradoxal, cet énoncé doit être accepté selon sa seule valeur sémiotique ou rhétorique et non pas pour sa valeur de vérité. Si cet énoncé est dans la rubrique de politique européenne d'un quotidien, il est tout à fait pertinent et  $\Pr(I_E/I_S) = 3/3$  et  $\Pr(I_N) = 1$ .

• *pertinence déontique* : c'est l'adéquation des énoncés (force illocutoire) aux rôles joués par les interlocuteurs dans le dialogue. La mesure est difficile car elle dépend de paramètres socioculturels. Le plus simple est certainement de la mesurer en *tout ou rien* sur l'échelle "acceptable"/"inacceptable".

$$P_e\{\alpha(E,D,t)\} = \{0, 1\}$$

Par exemple : soit un dialogue entre un client et un agent (guichetier dans un cinéma),

(a) « j'exige un billet » est une formule habituellement inacceptable car trop agressive,

(b) « je vous prie de bien avoir l'obligeance de me donner un billet » est une formule trop "ampoulée", donc également inacceptable

(c) « vous reste-t-il des places ? » est généralement plus acceptable.

### 8.3. Interaction et système de recherche d'information

Du point de vue recherche d'information, le dialogue homme-machine en langage naturel peut apporter une contribution dans la formulation des requêtes mais aussi dans le processus de recherche. Les apports du dialogue peuvent se situer aux trois moments classiques de la recherche d'information : formulation du besoin d'information, présentation des résultats de la recherche, évaluation de la pertinence de ces résultats [SAL 71, 83, 89],

- (a) l'utilisation du moteur de recherche peut être fondue dans un processus de guidage, pour en exploiter plus efficacement les ressources au moment de la formulation du besoin d'information. Dans certains moteurs ce besoin s'exprime à l'aide d'une expression booléenne de mots clés : il serait alors préférable de l'exprimer en langage proche de la langue naturelle (LN). *La formulation du besoin d'information* a pour but de négocier un sens pour l'expression de ce besoin, afin de permettre au système de se focaliser sur un ensemble de documents pertinents. Le passage d'un besoin d'information à son expression en des termes acceptables par le SRI est une tâche difficile car l'utilisateur "idéal" doit avoir des connaissances sur le système lui-même (modèles de recherche d'informations sous-jacents, informations sur le corpus disponible, connaissances du domaine couvert par le corpus, etc.). Inversement, le système doit connaître le contexte du besoin de l'utilisateur à savoir l'arrière-plan de connaissances de celui-ci [HOW 94]. La tâche du SRI est de construire un sens pour la requête le plus proche possible du besoin de l'utilisateur.
- (b) la présentation de la réponse fournie par le système, est généralement donnée par une liste de références aux documents de la base. Ici aussi il serait intéressant de justifier ces réponses ou de les commenter (notons cependant que certains systèmes donnent des extraits de documents permettant de contextualiser la réponse). *La présentation de la réponse* consiste pour le système à présenter l'ensemble des documents retrouvés à



l'utilisateur. Cette présentation constitue le lieu de la confrontation entre le sens que l'utilisateur cherche à exprimer, et le sens effectivement produit par le système. Il s'agit de confronter le problème d'information à résoudre et le contenu du corpus.

- (c) l'évaluation de la pertinence des documents retrouvés se fait selon les critères de pertinence propres au système. Ceci n'est pas toujours explicite pour l'utilisateur. *L'évaluation de la pertinence des documents* consiste à aider l'utilisateur à examiner la réponse obtenue en se déplaçant parmi les documents retrouvés [DEN 97]. Cette *navigation* permet à l'utilisateur d'évoluer librement avec ou sans aide parmi les documents, d'évaluer sa satisfaction, et de procéder à une reformulation (avec élargissement ou raffinement de la requête précédente par exemple) s'il n'est pas satisfait.

Ces trois grandes phases traduisent une utilisation itérative du SRI, dont l'objectif final est de satisfaire l'utilisateur. Evidemment les trois phases du cycle ne sont pas toujours séquentielles : l'utilisateur peut décider d'une stratégie entièrement exploratoire basée sur la navigation, et se laisser guider ainsi par les documents présentés tout au long de son parcours. C'est en ce sens qu'une meilleure intégration à un processus de dialogue peut être bénéfique [SCH 97].

### ***Que peut apporter le dialogue ?***

L'approche dialogique se distingue de la navigation dans la mesure où l'on formule directement ce que l'on désire : l'accès est plus immédiat si l'on sait formuler sa demande [ODD 77]. Le problème n'est plus de se repérer mais de respecter et de maîtriser le code de communication (mots, énoncés, conventions linguistiques, etc.). On peut supposer qu'après un rapide apprentissage, ce sera le cas pour un dialogue en langue naturelle (ou un sous-langage spécialisé) [SCH 97]. On attend donc de ce dernier, un allègement de la charge cognitive par une prise en charge de l'interaction et de la coopération [ROU 99] afin de :

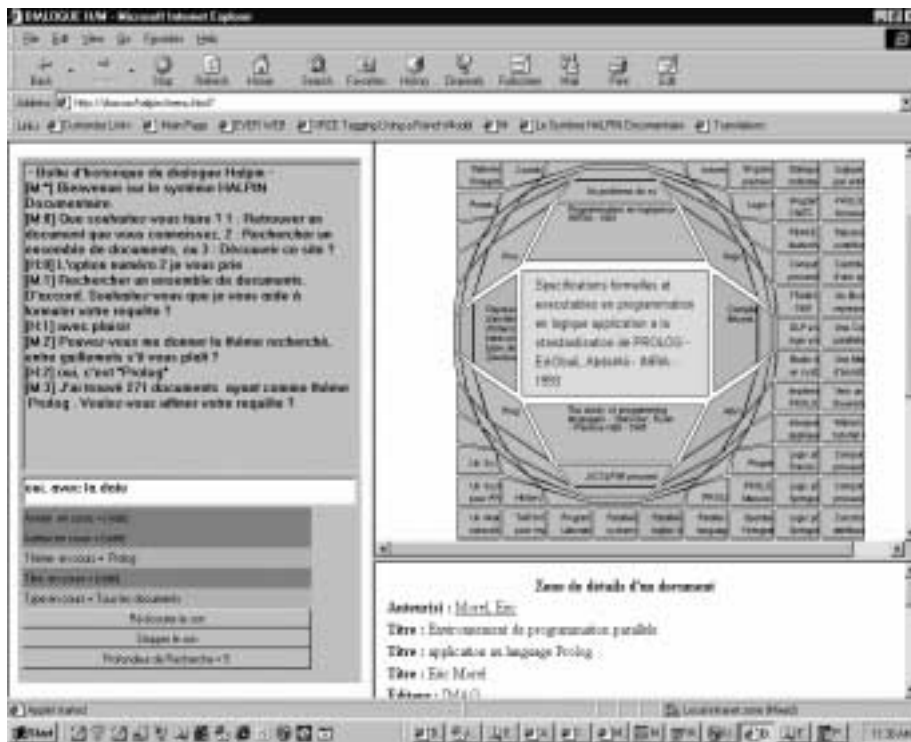
- rendre l'interaction plus souple par l'usage du langage naturel et des représentations graphiques,
- permettre un guidage efficace des utilisateurs par focalisation sur un thème ou un sous-thème donné, par élargissement ou affinement, thématique,
- permettre la formulation de requêtes imprécises : un nom d'auteur incomplet, une date approximative, etc. (CHE 97a, 97b),
- permettre l'usage de procédés tels que la reformulation, l'auto-correction, la clarification, etc. qui allègent la charge cognitive de l'utilisateur dans la mesure où il peut continuer à réfléchir tout en formulant sa demande,

– permettre un retour simple et direct aux documents déjà visités qui allège la charge cognitive de l'utilisateur en remplaçant un processus de rappel mémoriel par un processus de « dialogue ».

En Dialogue Homme-Machine (DHM), la robustesse des modèles de dialogue ainsi que la confrontation des prototypes à une utilisation réelle (hors laboratoire) demeurent des éléments importants pour valider toute nouvelle approche du problème. En effet, un système de DHM en LN doit pouvoir analyser et générer des énoncés aussi bien à propos de la tâche elle-même que du dialogue à propos de la tâche. En d'autres termes, un système sera jugé non seulement sur ses capacités dialogiques intrinsèques, mais également sur ses aptitudes à donner du sens à ce que dit son interlocuteur.

#### **8.4. Accès aux informations : le système multimodal HALPIN**

Le système HALPIN (Hyperdialogue avec un Agent en Langage Proche de l'Interaction Naturelle), que nous avons développé avec [ROU 98a ; 98b ; 98c], vise un meilleur accès aux informations par une interaction multimodale. La multimodalité est ainsi entendue à travers les modes d'entrée et de sortie des informations, en entrée, la parole, le texte écrit, les désignations graphiques (menus, boutons), et en sortie, la parole, le texte, le graphique en « œil de poisson » (voir fig. 8.3.). HALPIN utilise un niveau de dialogue multimodal ayant une composante en langage naturel écrit et/ou parlé. Il a été conçu dans un premier temps pour la recherche de documents *structurés*, puis, dans un deuxième temps, pour la recherche d'informations sur le Web.



**Fig. 8.3.** Interface du système documentaire HALPIN intégré dans un navigateur Microsoft. La fenêtre utile est divisée en quatre zones : la zone en haut à gauche contient l'historique du dialogue, celle du bas à gauche contient l'interface langagière d'entrée, la zone en haut à droite contient la présentation des documents en « œil de poisson » [VER 97] dans lesquels il est possible de naviguer, et la zone du bas à droite, les documents sélectionnés par navigation ou par dialogue.

La multimodalité est aussi entendue dans HALPIN du point de vue des modalités d'accès aux informations ; il utilise les deux modalités de recherche d'information : *par navigation, et par dialogue*. Vu par l'utilisateur, le système 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 joue, d'autre part, le rôle d'une bibliothèque en libre service où il est permis de butiner soi-même. HALPIN n'a pas de compétence sur le domaine et ne vise pas à en acquérir, bien que de telles connaissances soient certainement utiles dans sa tâche de recherche. 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. Le système HALPIN est donc essentiellement un *assistant dans la recherche immédiate et dans la collecte* : il permet à un utilisateur du Web de

dialoguer en langue naturelle avec une machine afin d'effectuer une recherche documentaire. Les hypothèses de travail de conception sont :

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

#### **8.4.1. Conception du système**

La méthode de conception d'HALPIN a suivi un cycle de développement en spirale : après analyse de l'usage et des fonctionnalités<sup>4</sup> nous avons développé une maquette que nous avons améliorée au fur et à mesure des tests. On remarque, en effet, que l'étude de la communication homme-machine ne peut se fonder uniquement sur le modèle de communication homme/homme même finalisé, car bien des phénomènes différents surviennent dans l'interaction homme-machine ; on se trouve ainsi en présence d'un problème circulaire [SIR 89] : celui de disposer de données pour réaliser un système qui n'a pas encore produit ces données. Ces insuffisances des connaissances et des modèles actuels constituent l'un des obstacles majeurs des recherches visant à la conception et à la mise en œuvre d'interfaces de dialogue en langue naturelle. Les expériences de type magicien d'Oz [DAH 93], par exemple - où un compère humain simule les comportements de la machine à l'insu des usagers - demandent beaucoup d'effort d'organisation et ne sont jamais à l'abri de biais expérimentaux. La méthode que nous avons choisie est donc la suivante :

- développer un système initial puis l'améliorer progressivement au vu des résultats que l'on obtient en utilisation réelle,
- en recueillant des corpus via le Web [MEI 98] où le DHM est réel (et crédible) avec un temps de réponse instantané (sauf lenteur locale du réseau Internet).

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

- 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,

---

<sup>4</sup> Cette analyse, non décrite ici, a consisté en une analyse d'usage auprès d'étudiants à la médiathèque de l'IMAG à Grenoble. Une série d'étudiants a été observée en situation et questionnée par une méthode d'enquête.

– quel est le système linguistique utilisé ? en termes de concepts.

Dans la phase préliminaire nous avons orienté le système HALPIN vers la collecte de connaissances sous forme d'un jeu interactif (diffusé sur Internet) en utilisant la technique du *jeu des questions* (il s'agit de faire deviner à quelqu'un, un objet que l'on a choisi mentalement, en ne répondant à ses questions que par oui ou par non). Cette technique permet généralement de repérer l'arbre des concepts que se représente le sujet. Après un recueil d'un millier de dialogues environ, nous avons pu dégager cet arbre de concepts et effectuer des analyses de dialogue et des analyses linguistiques (thématique, sémantique, syntaxique). Il en ressort principalement que :

– les concepts d'ouverture et de clôture de dialogue (concepts dits de *dialogue*) sont peu présents dans ce corpus. Ce phénomène est sans doute dû au fait que le dialogue est à l'initiative de l'utilisateur. 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.

– les concepts de clarification sont marqués par des questions.

– le concept de thème (concept dit du *domaine*) est souvent marqué par un substantif : « quel est le *sujet* du livre ? », « quels sont les *mots-clés* ? », « donne moi le *thème* du bouquin », mais aussi des phrases du type « de quoi ça parle ? » où ce n'est que le verbe qui peut fournir une information sémantique.

– les autres concepts du domaine (auteur, date, éditeur, type de document, etc.) sont également marqués par des substantifs ou des formules toutes faites.

– les concepts du domaine sont ordonnés dans la tâche : la plupart du temps le sujet fait référence au thème, puis à l'auteur et aux dates de parution. Il ne mentionne que rarement l'éditeur.

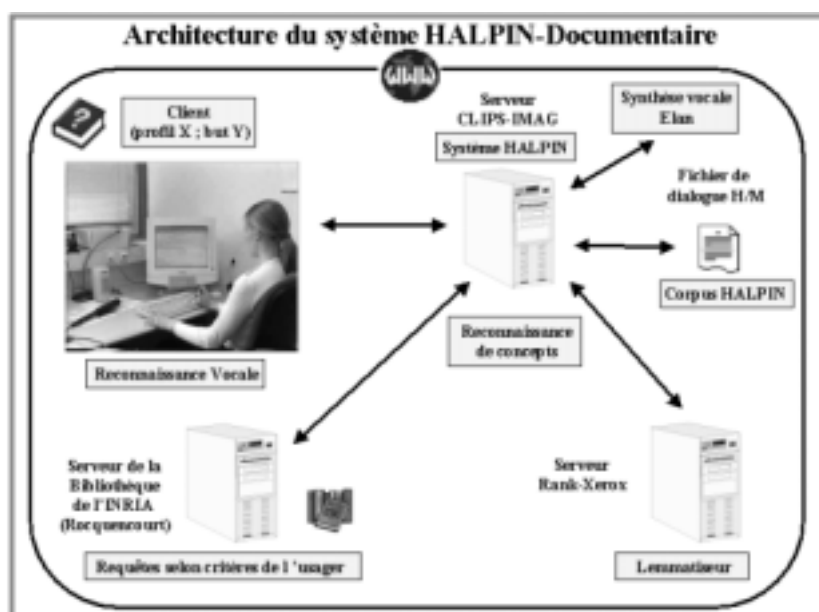
– en ce qui concerne le concept de confirmation (concept dit *linguistique*), on retrouve dans 79,63% des cas la lexie « oui ». Notons par exemple que 28 % des alternatives au « oui » comportent la locution « pourquoi pas » pour exprimer une approbation. Pour le concept d'infirmité, le vocabulaire se limite à trois termes principaux : « non » (66 occurrences, soit 13,58% de présence dans le corpus), puis « ne » et « pas » (33 et 37 occurrences, soit 6,79 et 7,61%).

Un total de 162 verbes les plus fréquents, a été identifié sur l'ensemble du corpus enregistré. On note bien entendu une importante présence des verbes d'états : respectivement 24,07 et 87,45% pour les auxiliaires « être » et « avoir ». On retrouve également les verbes de la tâche, comme chercher, rechercher, connaître, trouver, etc. Plus généralement, en comparant ces résultats avec d'autres corpus de dialogues Homme-Machine, on peut confirmer que le vocabulaire employé est souvent en étroite relation avec la tâche à réaliser [FRE 93a].

### 8.4.2. Réalisation du système HALPIN

Le système réalisé s'appuie sur un ensemble de modules implémentés sur différents serveurs distants communicant sous le protocole IP (on utilise des logiciels du commerce, des logiciels en libre service et des logiciels développés pour HALPIN). L'application finale tourne dans un navigateur sous environnement Unix, PC, Mac, ou tout autre système supportant un navigateur gérant les applets Java. La figure 4 présente l'architecture générale du système HALPIN-documentaire. Les modules sont :

- *La reconnaissance de la parole* : est fournie par le système ViaVoice d'IBM (et aussi par [AKB 98]). Ce module (sous licence) doit être installé sur la machine cliente. Il délivre en sortie une chaîne de caractères que l'utilisateur peut encore corriger à son gré avant de la valider. S'il ne dispose pas de ce logiciel, il peut entrer toutes les requêtes au clavier.



**Fig. 8.4.** : Architecture du système HALPIN : la requête est collectée sur un poste de travail local puis envoyée au serveur de dialogue Halpin qui utilise des ressources propres (reconnaissance vocale, gestion du dialogue et synthèse vocale) et des ressources distantes (lemmatiseur et base de données documentaire ou Web). La réponse (document trouvé ou acte dialogique suivant) est renvoyée à l'utilisateur sur son poste de travail.

– *L'analyseur morphologique* : est fourni par le serveur du Centre de Recherches de XEROX [GAU 97] auquel on accède de manière distante. Ce module délivre la forme lemmatisée de chacun des mots de la source d'entrée.

– *La compréhension par concepts* : est un module développé dans le cadre d'HALPIN [ROU 98a]. Il fournit une structure de traits à partir d'une analyse conceptuelle de la source lemmatisée.

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

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

– *La synthèse vocale* : est le système de la société française Elan informatique. Ce module fonctionne en serveur. Contrairement à la reconnaissance, l'utilisateur n'a pas besoin de l'installer sur sa propre machine.

– *L'interface graphique* tourne sous un navigateur comme Internet Explorer ou Netscape.

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

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

#### 8.4.2.1. Des énoncés aux concepts

La méthode de compréhension *par concepts* est une méthode *top-down* bien connue, elle a l'avantage d'être robuste pour des langages finalisés contraints par la tâche car elle permet de ne pas analyser totalement l'énoncé (cela permet de tolérer les erreurs d'insertion introduites par le système de reconnaissance, et certaines erreurs de substitution ou de délétion).

La méthode consiste à présupposer une intention communicative et à guider l'analyse par recherche de segments-clés ou de mots-clés. Par exemple avec le but présupposé DEMANDER(DOCUMENT), l'énoncé "Donne-moi un ouvrage de Baudelaire" sera analysé comme : Donne-moi = DEMANDER, ouvrage = DOCUMENT, Baudelaire = AUTEUR, les autres mots de l'énoncé n'étant pas consommés dans un premier temps par l'analyse. La structure de traits (attributs/valeurs) issue de l'analyse sera donc :

U(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

U(DEMANDER(DOCUMENT(THEME(Baudelaire))))). On voit donc toute l'importance de :

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

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

- Recherche d'expressions typées à l'aide d'un dictionnaire, pour construire des hypothèses de structures de traits
- Recherche de marqueurs grammaticaux (prépositions, négations, etc.) pouvant compléter ou orienter la structure de traits (comme l'exemple ci-dessus où *sur* oriente plutôt vers le trait *thème* que vers le trait *auteur*).

Cette méthode en deux étapes, constitue la base du processus de compréhension que nous détaillons maintenant.

#### 8.4.2.2. Des concepts aux actes

Au cours de l'étape précédente on a collecté dans l'énoncé tous les concepts possibles. Il s'agit maintenant de les filtrer et de les assembler pour former une structure de traits. On utilise pour cela une grammaire sémantique gauche-droite à rattachement incrémental. Il faut ensuite les interpréter vis-à-vis des actes de langage.

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 :

U(CONFIRMER)+U(REMERCIER)+U(DEMANDER(DOCUMENT(AUTEUR)))

Ces trois actes ont été rattachés à l'historique (confirmer et remercier) et à la situation courante (demander) car les deux premiers satisfont les actes précédents et le dernier non.

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



### 8.4.2.3. La gestion du dialogue

Le modèle de dialogue que nous avons choisi pour HALPIN, est un modèle de dialogue à stratégie dynamique. Le modèle utilise trois stratégies<sup>5</sup> :

– *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.

Pour la mise en œuvre du contrôleur de dialogue, nous nous sommes inspirés du modèle COR (CONversational Roles) de [STE 95], [BAT 95]. Ce modèle propose des séquences de dialogues *idéales* et *alternatives*. Par exemple, un dialogue entre E (qui demande de l'information) et D (qui en fournit) est formalisé comme suit :

Demande (E,D) => promesse (D, E) + information (D,E) + test-satisfaction (E,D)

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

U(DEMANDER(y))/[COOPERATIF] => M(PROMETTRE • INFORMER(y) • JUSTIFIER(y) • SUGGERER(z/BUT))

U(DEMANDER(y))/[REACTIF] => M(PROMETTRE • INFORMER(y))

U(ORDONNER(y))/[COOPERATIF] => M(PROMETTRE • EXECUTER(y) • JUSTIFIER(y) • SUGGERER(z/BUT))

U(ORDONNER(y))/[REACTIF] => M(PROMETTRE • EXECUTER(y))

---

<sup>5</sup> 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.

Etc.

Les notations [ ] signifiant en "contexte de", {} présence facultative et • est l'opérateur de séquence.

Ce qui se lit : la réponse à un acte de U par M est conditionnée par le niveau d'assistance de U. Dans le cas d'une demande en contexte coopératif, la réponse sera constituée d'une promesse (mise en attente), d'une information sur l'objet de la demande et éventuellement suivie d'une justification et d'une suggestion coopérative pour le tour de parole suivant.

Un modèle fondé sur la pertinence serait facile à implémenter avec ce formalisme [FOU 01] en ajoutant un niveau de traitement des attentes. On définit l'ensemble des ententes d'un acte par l'ensemble des réactions possibles à cet acte. Par exemple à la question « avez-vous l'heure ? », selon les circonstances ou le contexte cette demande peut-être interprétée de différentes façons : comme une invitation au dialogue, comme une demande de renseignements, comme un reproche, etc. Parmi celles-ci il en existe de plus pertinentes que d'autres. Un modèle plus fin que le précédent peut donc se formuler par :

Acte/[contexte] => {attentes} => Action/[maxP(attentes)]  
 où P(attentes) est la probabilité d'une attente dans un contexte donné. Par exemple dans l'un des cas explicité ci-dessus :  
 $U(\text{DEMANDER}(y))/[\text{COOPERATIF}] \Rightarrow \{\text{REPONSE}(y), \text{MISE-EN-CAUSE}(\text{DEMANDER}), \text{CLARIFICATION}(y)\}$   
 Si  $P(\text{REPONSE}) > P(\text{MISE-EN-CAUSE})$  ou  $P(\text{CLARIFICATION})$   
 Alors  $M(\text{PROMETTRE} \cdot \text{INFORMER}(y) \cdot \text{JUSTIFIER}(y) \cdot \text{SUGGERER}(z/\text{BUT}))$   
 Si  $P(\text{MISE-EN-CAUSE}) > P(\text{REPONSE})$  ou  $P(\text{CLARIFICATION})$   
 Alors  $M(\text{PROMETTRE} \cdot \text{SUGGERER}(z/\text{BUT}))$   
 Si  $P(\text{CLARIFICATION}) > P(\text{MISE-EN-CAUSE})$  ou  $P(\text{REPONSE})$   
 Alors  $M(\text{PROMETTRE} \cdot \text{DEMANDER}(y))$

Dans les deux cas, formulation simple ou plus fine, le modèle de dialogue peut être implémenté facilement par l'algorithme ci-après :

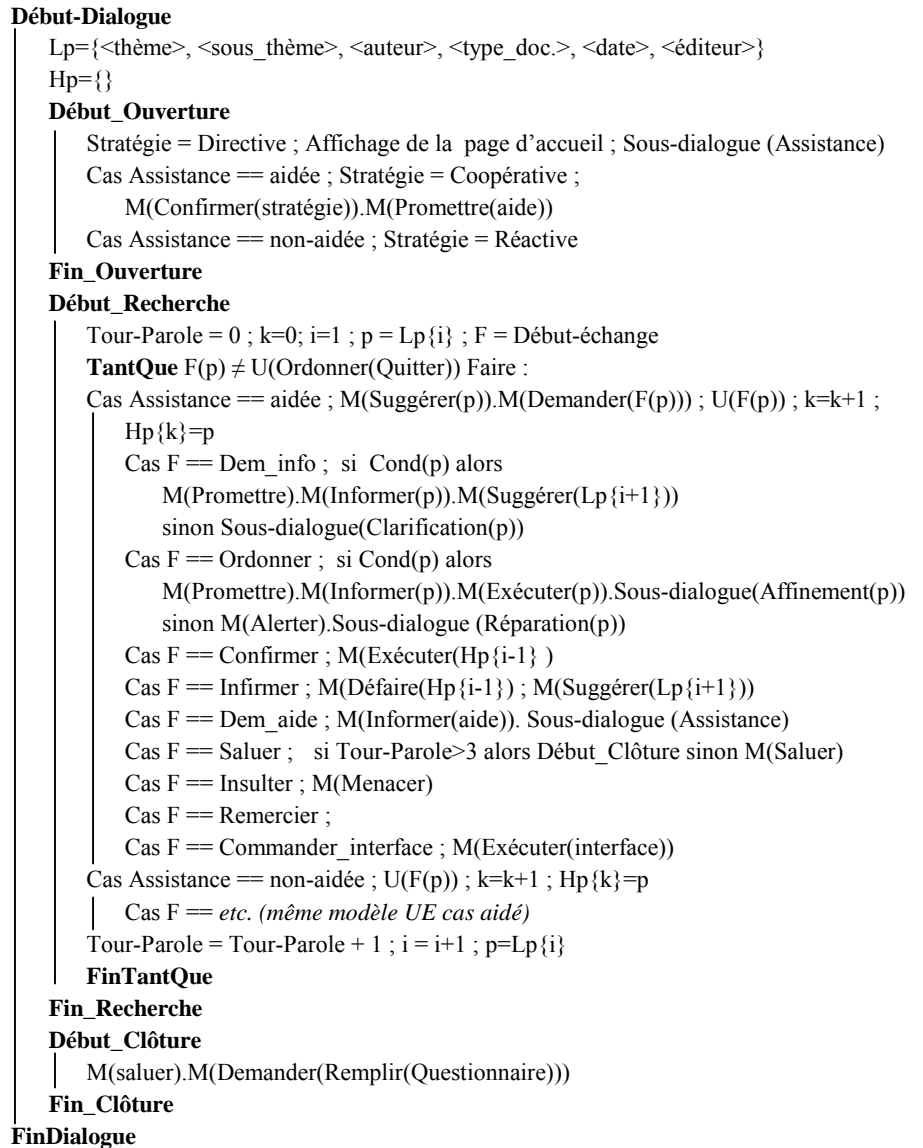
Légende :

F(p) acte de langage dont la force est F et le contenu propositionnel est p.

Lp liste des concepts de la tâche sur lesquels peuvent porter les énoncés du dialogue.

Hp historique du dialogue (suite temporelle des énoncés propositionnels).

Cond(p) ensemble des conditions que doit satisfaire p (définies par Vanderveken dans sa logique illocutoire, en résumé : conditions préparatoires, conditions de vérité, conditions d'arrière-plan, conditions de sincérité).



## 8.5. Résultats

Les résultats que nous exposons ici sont plus d'ordre qualitatif que quantitatif. Nous exposons tout d'abord des résultats portant sur l'évaluation du dialogue vs la

navigation, puis nous présentons les tendances qui se dégagent quant à la satisfaction des usagers lors de l'utilisation d'HALPIN sur le Web.

### 8.5.1. Résultats portant sur le dialogue vs. navigation

L'objectif est d'évaluer l'aide apportée par le dialogue et la navigation. L'hypothèse *a priori* est que le dialogue doit permettre une recherche plus efficace et qu'en second lieu la classification thématique des documents retournés par le système, doit améliorer la navigation. Aucune hypothèse n'est faite sur la supériorité de l'une ou de l'autre méthode d'accès. Le test d'évaluation est fait auprès de 10 sujets soumis à 4 tests de recherche documentaire, sans limite de temps. Le but de l'expérimentation est de comprendre dans quelles circonstances l'utilisateur utilise plus volontiers un dialogue qu'une navigation, ou l'inverse.

**1<sup>er</sup> test** : a pour but de tester la recherche d'une information particulière d'un document. La consigne donnée aux sujets est : « Rechercher le titre d'un *rapport* (type de document) écrit par *Pierrel* en 1992 ». Le temps moyen pour effectuer cet exercice a été de 6 min. 2 sujets ont utilisé uniquement le dialogue, les 8 autres ont d'abord dialogué puis ont navigué. 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<sup>ème</sup> test** : a pour but de tester la recherche d'un document. La consigne donnée aux sujets est : « Recherchez un document ayant pour thème *parole* et édité par *l'INRIA* entre 1990 et 1993. ». Le temps moyen pour effectuer cet exercice a été de 9 min. Notons qu'il n'était pas possible de soumettre une requête sur le critère *éditeur*. Les sujets ont tous utilisé le dialogue pour affiner leurs requêtes selon le critère année.

A l'issue de ces 2 tests, on peut remarquer que l'usage du dialogue ne dépend pas du but de l'utilisateur (recherche précise ou imprécise de document). L'usage du dialogue se fait préférentiellement à la navigation, celle-ci n'étant utilisée qu'en second lieu ou par substitution en cas d'échec.

**3<sup>ème</sup> test** : a pour but de tester la recherche de plusieurs documents. Il se décompose en deux tâches relativement proches ; avec et sans utilisation de la représentation en « œil de poisson » (notée OP ci-après) :

- « Rechercher le nombre de documents écrits en anglais ayant pour thème *dialogue* » (avec OP). Là encore, il n'était pas possible d'effectuer une requête prenant en compte la langue du document recherché comme critère.
- « Rechercher le nombre de documents écrits en anglais ayant pour thème *discours* » (sans OP). C'était donc le même exercice mais avec une liste

d'hyperliens à la place de l'OP. Le temps moyen pour réaliser cet exercice diminue sensiblement par rapport à l'exercice précédent (2 minutes contre 4 minutes). Les sujets ont expliqué qu'il était plus facile de dénombrer les documents dans la liste et qu'ainsi on pouvait immédiatement identifier leurs.

Ce test montre qu'une représentation en OP n'est pas toujours pertinente pour la navigation : il faut que la nature des données s'y prête.

**4<sup>ème</sup> test** : a pour but de tester la recherche indirecte d'une information à propos d'un document. La consigne donnée aux sujets est : « Rechercher, dans les documents ayant pour thème la *programmation par objets*, 5 sous-thèmes qui lui sont liés ». 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. Cet exercice montre les limites cognitives dans l'utilisation du dialogue : l'utilisateur ne pense pas la machine capable de dialoguer de manière pertinente.

Le module intégré de synthèse vocale est probablement un fort incitateur de maintien du dialogue qui certainement biaise un peu les résultats. Nous devons encore analyser jusqu'à quel point le dialogue en langage naturel peut se substituer à une interaction Homme-Machine traditionnelle (clavier/souris), mais nous avons montré que, toute chose étant égale par ailleurs (sur une base de données identique), une interaction multimodale est une réponse efficace aux problèmes de désorientation, de surcharge cognitive, et d'évaluation de la pertinence des réponses du système. Les exemples de dialogues réels, figures 5 et 6, recueillis avec le système HALPIN montrent comment la machine se comporte dans le dialogue.

<p>11:59:18:M=Bonjour. Que souhaitez-vous faire ? Retrouver un document que vous connaissez, Rechercher un ensemble de documents ou Découvrir ce site ?</p> <p>12:00:03:U= un ensemble de documents</p> <p>12:00:15:M=Rechercher un ensemble de documents. D'accord. Souhaitez-vous que je vous aide à formuler votre requête ?</p> <p>12:00:30:U=oui</p> <p>12:00:37:M=Pouvez-vous me donner le thème recherché s'il vous plaît ?</p> <p>12:01:11:U=karaoke</p> <p>12:02:18:M=Je n'ai pas trouvé de document ayant comme thème karaoke. Voulez-vous modifier votre requête ?</p> <p>12:03:51:U=je veux chercher avec synthèse de la parole</p> <p>12:04:09:M=J'ai trouvé 398 documents ayant comme thème synthèse de la parole. Voulez-vous affiner votre requête ?</p> <p>12:05:28:H=je veux l'année 1998</p>
---

12:06:01:M=J'ai trouvé 12 documents de 1998 ayant comme thème synthèse. Voulez-vous affiner votre requête ?  
12:06:46:U=je ne vois rien  
12:06:58:M=Okay j'augmente la taille de la police.  
12:07:09:U=efface tout  
12:07:24:M=D'accord, tous les critères sont vides à présent. Faites un choix parmi les critères : Auteur, Titre, Thème, Année, ou Type de document.  
12:07:35:U=auteur  
12:07:49:M=Vous pouvez soit m'indiquer le nom de l'auteur, avec une majuscule, soit effacer l'auteur en cours. Ou bien dites Non pour passer à autre chose.  
12:08:08:U=je veux Astrom  
12:08:22:M=J'ai trouvé 5 documents avec comme auteur Astrom. Consultez-les sur la page de droite. Est-ce que cela vous convient ?  
12:10:20:U=oui  
12:10:30:M=Très bien, je suis heureux d'avoir pu vous aider dans votre recherche.

**Fig. 8. 5.** Extrait d'un dialogue du corpus HALPIN-Documentaire

Dans l'exemple de la figure 8.5 on observe bien entendu des tours de parole classiques à propos de la recherche d'informations, mais on en note également d'autres concernant l'interface, comme par exemple l'intervention 12:06:46:U.

15:57:58:M=J'ai trouvé 1204 documents ayant comme thème intelligence artificielle. Voulez-vous raffiner votre requête ?  
15:58:18:U=oui, il y en a beaucoup trop  
15:59:07:M=Le thème actuel est intelligence artificielle. On pourrait y ajouter par exemple : représentation connaissance ; langage naturel ; résolution problème ; apprentissage ; acquisition connaissance. Pour ajouter un thème, donnez le entre guillemets, sinon répondez Non pour passer à autre chose.  
16:00:21:U=oui, voyons avec "langage naturel"  
16:00:51:M=J'ai trouvé 174 documents ayant comme thème intelligence artificielle langage naturel. Voulez-vous raffiner votre requête ?  
16:01:04:U=ben oui  
16:01:23:M=Afin de raffiner, vous devez choisir parmi les critères : Auteur, Titre, Thème, Année, ou Type de document.  
16:01:58:U=l'année 98  
16:02:17:M=J'ai trouvé 4 documents de 1998 ayant comme thème intelligence artificielle langage naturel. Consultez-les sur la page de droite. Est-ce que cela vous convient ?  
16:02:42:U=ouais, ça a l'air bon  
16:02:51:M=Très bien, je suis heureux d'avoir pu vous aider dans votre recherche.

**Fig. 8.6.** De 1204 à 4 réponses : affinement des thèmes. On a typiquement un cas où le système HALPIN aide l'utilisateur de manière coopérative, en tenant compte de son but. On passe donc de 1204 réponses à 4 réponses par raffinement de la recherche

### 8.5.2. Résultats sur la satisfaction des utilisateurs

Les usagers qui se sont connectés au système HALPIN avaient la possibilité de répondre à un questionnaire d'évaluation, en ligne. Globalement, l'interaction dialoguée avec le système HALPIN est considérée, par les 19 personnes interrogées, comme bonne à 55,6%, mauvaise à 33,3% et très mauvaise à 11,1%. A la question « *Le dialogue avec la machine vous a-t-il orienté dans votre recherche ?* », 23,5% ont répondu « Pas assez », 47,1% « Suffisamment », et 29,1% « Un peu ».

Les usagers déclarent à 58,7% que le dialogue leur a permis de trouver ce qu'ils cherchaient (*en partie* : 23,5% ; *un peu* : 17,6% ; *totale*ment : 17,6 %) contre 41,3% n'ayant *pas du tout* retrouvé l'information recherchée. Le dialogue est perçu comme robuste pour la moitié des personnes interrogées, et 56,3% ont apprécié l'aide apportée par la machine lors des interactions dialoguées.

La synthèse vocale utilisée est considérée comme « *très bonne* » pour 21,1% des utilisateurs, « *bonne* » pour 47,4% et « *mauvaise* » pour 10,5% (21% de sans opinion). Elle leur semble très intelligible (42,1 %) et moyennement naturelle (36,8%). Les avis à propos de l'utilité du système, comparé à un formulaire traditionnel du Web, sont que dans l'ensemble, 41% des personnes trouvent ce système moins utile qu'une interaction traditionnelle, 12% la trouvent similaire et 47% la trouvent meilleure.

Les temps de réponse du système, quant à eux, sont jugés comme étant *rapide* à 15,8%, *acceptable* à 36,8% et *trop long* à 47,4%. La majorité des personnes ayant répondu sont des hommes (79%). La plupart sont des informaticiens (58%) (et 12% de bibliothécaire). Notons enfin que les salariés sont représentés à 42%, les enseignants et/ou chercheurs à 32%, et les étudiants à 26%.

## 8.6. Conclusion

La mise à disposition de grands volumes d'information complexes et hétérogènes oblige à définir des moyens d'accès efficaces et précis. De plus en plus, la qualité de l'interaction avec le système de recherche d'information jouera un rôle important. L'utilisateur est plus efficace lorsqu'il peut composer avec ses capacités naturelles : les cinq sens et le langage. La dimension vocale (reconnaissance de la parole,

dialogue et synthèse sonore) s'avère précieuse comme support d'échange entre utilisateur et machine. Elle apporte en effet trois atouts majeurs :

- la possibilité d'utiliser la langue naturelle comme moyen d'interaction spontané, notamment en vue d'une plus grande richesse (et souplesse) d'expression,
- la conservation de l'habitude de la construction de la pensée à travers l'argumentation par le dialogue, notamment en maintenant toutes les ressources cognitives concentrées sur la tâche plutôt que sur le mode d'interaction,
- la possibilité de varier les modes d'interaction pour une meilleure efficacité dans la tâche (navigation opportuniste vs. dialogue coopératif finalisé).

Le système Halpin représente une première approche vers cette voie : utilisation du dialogue multimodal en lui conservant la possibilité de navigation classique par liens hypertextuels. Il va de soi que des améliorations devront encore être apportées quant à la robustesse du système et à ses capacités de compréhension. Une meilleure gestion des procédés anaphoriques<sup>6</sup>, ainsi qu'une gestion des ellipses<sup>7</sup> ou des tropes<sup>8</sup> est à prendre en compte, car ils assurent la concision et la souplesse du dialogue en langage naturel. Mais plus que sur les problèmes d'ordre technique, c'est en matière de pertinence du modèle de l'utilisateur qu'il faut encore travailler. D'ores et déjà ce système illustre bien l'utilité d'une composante d'interaction multimodale pour la recherche d'information.

**Remerciements** : le présent travail a été possible grâce aux nombreux échanges sur le plan théorique avec D. Vernant et D. Vanderveken. La réalisation informatique a été effectuée par J. Rouillard. Enfin le support financier a été assuré par la Région Rhône-Alpes (projet ORION), par le Consulat de France à Québec et le Ministère des Affaires Internationales du Québec (projet Dialogue pour le Web).

## 8.7. Bibliographie

- [AKB 98] AKBAR M., CAELEN J., "Parole et traduction automatique : le module de reconnaissance RAPHAEL", Proceedings COLING-ACL'98, Montréal (Québec), p. 36-40, August 1998.

---

<sup>6</sup> L'anaphore est un mot renvoyant à un membre de phrase déjà énoncé.

<sup>7</sup> L'ellipse est l'omission d'un ou de plusieurs mots d'un énoncé sans altération du sens de l'énoncé.

<sup>8</sup> Un trope est une figure rhétorique consistant à employer un mot ou une expression en les détournant de leur sens propre.



- [ALL 84] ALLEN J., Towards a general theory of action and time. *Artificial Intelligence*, 23, p. 123-154, 1984.
- [ALL 80] ALLEN J., PERRAULT C.R. Analyzing intentions in utterances. *Artificial Intelligence*, 15(3), p. 143-178, 1980.
- [ARM 85] ARMENGAUD F., La pragmatique, coll. Que sais-je?, PUF éd., 1985.
- [AUS 62] AUSTIN J.L., How to do Things with Words. Oxford University Press, 1962.  
Version française : Quand dire c'est faire. Seuil, Paris, 1970.
- [BIL 92] BILANGE E., Dialogue personne-machine, *Hermès éd.*, Paris, 1992.
- [BRA 87] BRATMAN M., Intentions, plans and practical reason. Cambridge MIT Press, 1987.
- [CAE 95] CAELEN, J., Vers une logique dialogique. Séminaire international de pragmatique, Jérusalem, 1995.
- [CAR 90] CARBERRY S., Plan Recognition in Natural Language Dialogue, MIT Press, Cambridge, 1990.
- [CAT 95] CATINIS L., CAELEN J., Analyse du comportement multimodal de l'utilisateur humain dans une tâche de dessin. Actes du congrès IHM'95, Cepadues éd., Toulouse, p. 123-130, 1995
- [CHE 97a] CHEVALLET J.P., M.-F. BRUANDET M.F., NIE J., "Impact de l'utilisation de multitermes sur la qualité des réponses d'un système de recherche d'informations à indexation automatique", Actes des Premières Journées du Chapitre Français de l'ISKO, Lille (France), octobre 1997.
- [CHE 97b] CHEVALLET J.P., NIE J.Y., "Intégration des Analyses du Français dans la Recherche d'Information", Actes de la Conférence Recherche d'Information Assistée par Ordinateur (RIAO) 97, Montréal (Canada), p. 761-772, juin 1997.
- [COH 78] COHEN P.R., On knowing what to say : Planning speech acts. Ph.D. Thesis, Technical Report n°118, Department of Computer Science, University of Toronto, January 1978.
- [COH 79] COHEN P.R., PERRAULT C.R., Elements of a Plan-Based Theory of Speech Acts. *Cognitive Science* 3, p. 177-212, 1979.
- [COH 90] COHEN P.R., LEVESQUE H.J., Rational interaction as the basis for communication. in P.R. Cohen, J.L. Morgan, M.E. Pollack (eds.), *Intentions in communication*. Cambridge MIT Press, 1990.
- [DAH 93] DAHLBACK N., JONSSON A., AHRENBERG L., "Wizard of Oz studies – why and how", *Knowledge-based systems*, December Issue, vol. VI.4, p. 258-266, 1993.
- [DEN 97] DENOS N., Modélisation de la pertinence en recherche d'information - modèle conceptuel, formalisation et application, thèse d'université, Grenoble, octobre 1997.

[ELAN] <http://www.elan.fr>

[FRE 93a] FRÉCHET A.L., Analyse linguistique d'un corpus de dialogue Homme-Machine (oral finalisé), thèse d'université, Université de la Sorbonne nouvelle, Paris III, 1993.

[FOU 01] FOUQUET Y., De l'étude de dialogues oraux dans une langue non maternelle. Journées RJC2001, Martigny, 2001

[FRE 93b] FRÉCHET A.L., CAELEN J., Cognitive Attitudes and Speech Acts in Situations of Man-Machine Communication. Selected Proceedings of the 5rd International Conference Work With Display Units (WWDU'92), Berlin, Germany, H. Luczak, A. Cakir and G. Cakir eds., Amsterdam, Elsevier, pp. 340-344, 1993.

[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'AUFELF-UREF, Avignon, Avril 1997.

[GOF 81] GOFFMAN E., Forms of talk. Oxford, Basil Blackwell ed., 1981.

[GRI 75] GRICE H.P., Logic and Conversation. in P. Cole and J.L. Morgan eds. *Syntax and Semantics*, vol. 3, Academic Press, p. 41-58, 1975.

[GRO 86] GROSZ B.J., SIDNER C.L., Attention, Intention in the Structure of Discourse, *Computational Linguistics* 12(3), p. 175-204, 1986.

[GRO 90] GROSZ B.J., SIDNER C.L., Plans for the discourse. in P.R. Cohen, J.L. Morgan, M.E. Pollack (eds.), *Intentions in communication*. Cambridge MIT Press, 1990.

[GUY 93] GUYOMARD M., NERZIC P., SIROUX J., Plans métaplans et dialogue. Actes de la 4ème école d'été sur le Traitement des langues naturelles. Lannion, 1993.

[HOW 94] HOWARD D. L., "Pertinence as reflected in personal constructs", *Journal of the American Society for Information Science*, vol. 45, n° 3, p. 172-185, 1994.

[HUL 00] HULSTIJN, J., Dialogue models for inquiry and transaction. Ph.D. thesis, University of Twente, The Netherlands, 2000

[LEH 97] LEHUEN J., Un modèle de dialogue dynamique et générique intégrant l'acquisition de sa compétence linguistique. Thèse de doctorat de l'université de Caen, 1997.

[LIT 87] LITMAN D.J., ALLEN J.F., A Plan Recognition Model for Subdialogues, *Cognitive Science*, 11, p. 163-200, 1987.

[LIT 90] LITMAN D.J., ALLEN J.F., Discourse processing and commonsense plans. in P.R. Cohen, J.L. Morgan, M.E. Pollack (eds.), *Intentions in communication*. Cambridge MIT Press, 1990.

[LUZ 89] LUZZATI D., *Recherches sur le dialogue Homme-Machine: modèles linguistiques et traitement automatique*. Thèse d'Etat, Paris III, 1989.

- [LUZ 95] LUZZATI D., Le dialogue verbal homme-machine. Masson éd., Paris, 1995.
- [MAU 01] MAUDET N., Modéliser les conventions des interactions langagières : la contribution des jeux de dialogue. Thèse de l'université Paul Sabatier, Toulouse, 2001
- [MEI 98] MEIYE J.P., ROUILLARD J., VAUFREDAZ D., Webcompletion - protocole de normes associatives sur Internet, Actes de l'Ecole thématique CNRS - Sémantique, Caen (Asnelles-sur-mer), Janvier 1998.
- [MOE 89] MOESCHLER J., Modélisation du dialogue, représentation de l'inférence argumentative. Hermès ed., 1989.
- [MOE 94] MOESCHLER J., REBOUL A., Dictionnaire encyclopédique de pragmatique, édition du Seuil, 1994.
- [NER 93] NERZIC P., Erreurs et échec dans le dialogue oral homme-machine: détection et réparation. *Thèse de doctorat de l'univeristé de Rennes*, 1993.
- [NIL 80] NILSSON H.J., Principles of artificial intelligence. Tioga ed., Palo Alto, CA, USA, 1980.
- [ODD 77] ODDY R., Information retrieval through Man Machine Dialogue, *Journal of Documentation*, 33, p. 1-44, 1977.
- [OZK 93] OZKAN N., CAELEN J., Towards an adaptative dialogue model, Actes du symposium "Interfaces des mondes réels et virtuels", Montpellier, pp. 77-84, 1993.
- [OZK 94] OZKAN N., Analyse du dialogue dans une perspective communicationnelle, thèse de 3ème cycle à l'Institut Polytechnique de Grenoble, 1994.
- [PER 80] PERRAULT C.R., ALLEN J.F., A plan based analysis on indirect speech acts. *American Journal of Computational Linguistics*, 6(3), p. 167-182, 1980
- [PIA 64] PIAGET J., Development and Learning. Piaget Rediscovered, Ripple R.E. and Roccastle V.N. ed., School of Education, Cornell University, Ithaca, New York, 1964.
- [ROU 85] ROULET E., L'articulation du discours en français contemporain. Peter Lang ed., Berne, 1985.
- [ROU 98a] ROUILLARD J., CAELEN J., Etude du dialogue Homme-Machine en langue naturelle sur le Web pour une recherche documentaire, Actes du Deuxième Colloque International sur l'Apprentissage Personne-Système, CAPS'98, Caen, Juillet 1998.
- [ROU 98b] ROUILLARD J., Hyperdialogue Homme-Machine sur le World Wide Web : Le système HALPIN, Colloque International Ergonomie et Informatique, ERGO'IA, Biarritz, Novembre 1998.
- [ROU 98c] ROUILLARD J., Contribution à l'étude du dialogue Homme-Machine à travers le Web : la personnalisation, Proceedings RECITAL'98, Le Mans, Septembre 1998.

- [SAL 71] SALTON G., The SMART Retrieval System: Experiments, in Automatic Document Processing, Prentice Hall Series in Automatic Computation, 1971.
- [SAL 83] SALTON G., MCGILL M.J., Introduction to Modern Information Retrieval, Mc Graw-Hill Computer Science Series, 1983.
- [SAL 89] SALTON G., Automatic Text Processing, Addison Wesley Series in Computer Science, 1989.
- [SEA 72] SEARLE J.R., Speech Acts. Cambridge University Press, UK., 1969. traduit de l'anglais "Actes de langage", Hermann éd., Paris, 1972.
- [SEA 83] SEARLE J.R., Intentionality. Cambridge University Press, UK., 1983.
- [SEA 85] SEARLE J.R., VANDERVEKEN D., Foundations in illocutionary logic. Cambridge University Press, 1985.
- [SAD 91] SADEK D., Attitudes mentales et interaction rationnelle : vers une théorie formelle de la communication. Thèse de doctorat de l'université de Rennes I (IRISA), 1991.
- [SAD 96] SADEK D., Le dialogue homme-machine : de l'ergonomie des interfaces à l'agent intelligent dialoguant. OFTA, ARAGO 18 Nouvelles interfaces homme-Machine, Lavoisier éd., Paris, p. 277-321, 1996.
- [SCH 97] SCHULTZ T., WESTPHAL M., WAIBEL A., "The GlobalPhone Project: Multilingual LVCSR with JANUS-3", Proceedings Multilingual Information Retrieval Dialogs: 2nd SQEL Workshop, p. 20-27, Plzen, Czech Republic, April 1997.
- [SHE 77] SHEGLOFF E.A., JEFFERSON G., SACKS H., The preference for self-correction in the organisation of repair in conversation. *Language*, 53, p. 361-382, 1977.
- [SID 85] SIDNER C.L., Plan parsing for intended response recognition in discourse. *Computational Intelligence*, 1 p. 1-10, 1985.
- [SIR 89] SIROUX J., GILLOUX M., GUYOMARD M., SORIN C., "Le dialogue homme-machine en langue naturelle : un défi ?", *Annales Télécommunication*, 44, n° 1-2, 1989.
- [SPE 86] SPERBER D., WILSON D., Relevance : communication and cognition. Basil Blackwell éd., Oxford, 1986.
- [STE 95] STEIN A., MAIER E., Structuring collaborative information-seeking dialogues, *Knowledge-Based Systems*, 8(2-3), Special Issue on Human-Computer Collaboration, p. 82-93, 1995.
- [STE 97] STEIN A., GULLA J. A., MULLER A., THIEL U., Conversational interaction for semantic access to multimedia information, in M.T. Maybury (Ed.), *Intelligent Multimedia Information Retrieval* (p. 399-421), Menlo Park, CA: AAAI/The MIT Press, 1997.
- [SUC 90] SUCHMAN L.S., Plans and Situated Actions, Cambridge University Press, 1987

- [VAN 90] VANDERVEKEN D., La logique illocutoire. Mandarga éd. Bruxelles, 1990.
- [VAN 91] VANDERVEKEN D., Meaning and speech acts, Vol. I & II, Cambridge University Press, 1991
- [VER 92] VERNANT D., Modèle projectif et structure actionnelle du dialogue informatif. in Du dialogue, Recherches sur la philosophie du langage, Vrin éd., Paris, n°14, p. 295-314, 1992
- [VER 97] VERNIER F., NIGAY L., "Représentations Multiples d'une Grande Quantité d'Informations", Actes IHM'97, Neuvièmes Journées sur l'ingénierie de l'Interaction Homme-Machine, Futuroscope de Poitiers, France, Cépaduès-éditions, p. 79-87, 1997.
- [VIL 92] VILNAT A., NICAUD L., Un système de dialogue homme-machine : STANDIA. Actes du séminaire Dialogue du GdR-PRC Communication homme-machine, Dourdan, 1992.
- [WIL 83] WILENSKY R., *Planning and Understanding. A computational approach to human reasoning*. Reading, Addison-Wesley Publishing Company, 1983.
- [ZAN 97] ZANELLO M.L., CAELEN J., BISSERET A., Une approche centrée tâche de la multimodalité. Actes du congrès IHM'96 , Cépaduès éd., 1997.