

Résumé de la thèse de Mme Djeradi Rachida

**Interface multimodale en communication homme machine.
Réalisation autour d'une application.**

Les activités de recherche intervenant dans les CPS (communication personne système), ont traité les problèmes d'interaction selon des approches tout d'abord purement informatique ou logico-mathématique, on a exclu l'utilisateur car on a supposé qu'il peut s'adapter à la machine. Depuis l'essor des interfaces graphiques et la diffusion des ordinateurs personnels, cette idée est complètement dépassée, on doit maintenant adapter la machine à l'utilisateur et partir de l'idée de pouvoir concevoir des systèmes utiles, utilisables, ergonomiques et intelligents.

L'interaction vocale entre l'homme et la machine est actuellement un domaine en pleine expansion, cependant la mise en place de ce type d'interface vocale, pose beaucoup de problèmes, surtout en ce qui concerne les utilisateurs et l'environnement externe. L'interface vocale doit être capable d'interagir avec n'importe quel utilisateur d'une façon robuste d'une part (c'est-à-dire qu'elle doit tenir compte par exemple de l'accent régional de l'utilisateur, des mauvaises prononciations, des hésitations, des faux départs et du bruit environnant) et d'autre part elle doit permettre à l'utilisateur d'arriver à son but sans blocage.

Dans le cadre de ce travail plusieurs réalisations d'interfaces ont été étudiées et mises en place. Nous avons tout d'abord élaboré une interface multimodale dans laquelle nous avons étudié la modalité parole et la modalité geste de désignation à l'aide de la souris pour remplir un document administratif.

Nous avons réalisé une autre application : Dans cette interface nous avons mis en place les étages de reconnaissance, de synthèse et un gestionnaire de dialogue entrant dans la communication entre l'homme et la machine. Ce dialogue tourne autour d'une application graphique. Dans cette étude, nous avons examiné les relations entre les sorties de l'étage de reconnaissance et les entrées du gestionnaire de dialogue d'une part et les sorties du gestionnaire de dialogue et du système de synthèse de parole d'autre part.

Une amélioration de ces travaux est l'utilisation d'un modèle de dialogue plus performant et plus complexe à fin de rendre l'interaction entre l'homme et la machine plus naturelle, ergonomique et plus agréable. Ça sera l'objet du point suivant

Dans la troisième partie de ce travail nous nous sommes intéressés à la modélisation d'un dialogue tournant autour de l'application « gestion de la scolarité ». Pour intégrer les représentations des connaissances nécessaires au fonctionnement de notre modèle, nous avons commencé par réaliser un système de base simple (question-réponse), puis nous avons fait une simulation de la tâche de recherche des notes assistée par ce système de dialogue. Cela nous a permis de saisir les représentations conceptuelles du sujet et d'y associer des réponses. Nous procédons ensuite au développement de notre système en utilisant une analyse morphologique de façon à regrouper des items en champs conceptuels et à vérifier qu'ils peuvent être des déclencheurs pertinents pour la tâche. En effet, il est possible d'incrémenter la base de connaissance du système de dialogue en lui ajoutant de nouvelles possibilités de compréhension contenant des formes nécessaires au dialogue.

Ce système est utilisé pour recueillir un corpus de dialogues Homme Machine réels. L'analyse de ce corpus nous a permis d'identifier les différentes composantes qui ont servi à la réalisation d'un modèle d'interaction ergonomique et intelligent.

*Enfin après ces expériences nous avons remarqué que l'élément central de ces systèmes de communication homme-machine reste le module d'interprétation et ceci quelque soit la modalité utilisée. Dès lors, nous nous sommes plus particulièrement intéressés à la mise au point d'un modèle comprenant à la fois la représentation des énoncés linguistiques et extra linguistiques c'est à dire capable de gérer le dialogue en **contexte**.*

Pour ce faire nous avons abordé les méthodes dynamiques d'analyse de discours en particulier la SDRT. Nous lui consacrons le paragraphe suivant. Enfin nous appliquons cette méthode à l'analyse de notre corpus dialogique.

I. INTRODUCTION

Les travaux de recherche sur la Communication Homme Machine visent à concevoir des interfaces ergonomiques et intelligentes. Ces systèmes doivent être capables d'interagir avec n'importe quel utilisateur d'une façon robuste, tenant compte d'une part, des capacités cognitives de celui-ci et d'autre part, lui permettre d'arriver à son but sans blocage. On voit donc que la modélisation de telles interfaces en communication homme-machine (CHM) s'inscrit dans un domaine pluridisciplinaire (linguistique, informatique, sciences cognitives ...).

Nous pouvons situer la CHM dans le cadre des relations entre trois entités : l'opérateur (l'utilisateur), la machine et l'environnement commun (l'application). Les échanges d'informations entre l'opérateur et la machine qui constituent la communication homme-machine s'effectuent dans le but de mener à terme une certaine tâche.

Par conséquent la machine doit donc disposer des représentations de l'opérateur (représentation orientée par la nature de la tâche à accomplir) et d'une stratégie permettant de suivre la réalisation de la tâche en coopération avec l'utilisateur, la CHM ne peut être réduite à un simple échange d'informations, elle doit en fait assurer la coordination entre deux processus intelligents qui se déroulent l'un dans le cerveau de l'utilisateur, l'autre dans la machine.

On mesure donc toute la difficulté, mais aussi tout l'intérêt de concevoir et de construire des systèmes "interfaces" dialoguant en langage naturel.

Le travail réalisé dans cette thèse comporte sept chapitres, le premier est consacré à l'étude de l'interaction entre les humains qui nous permettra d'aborder par la suite le domaine de la communication homme-machine.

L'objet du deuxième chapitre est l'étude et la réalisation d'une interface multimodale tournant autour d'une application " la gestion de la scolarité".

Dans le troisième chapitre, nous allons proposer une interface dans le cas d'une application graphique.

Le quatrième chapitre présente l'état de l'art sur les modèles de dialogues

Le cinquième sera consacré à un système d'interaction bilingue, évolutif et distribué. Notre démarche serait de pouvoir intégrer les représentations de nouvelles connaissances acquises au cours du dialogue.

Ce système sera utilisé pour recueillir un corpus de dialogues Homme Machine réels dans les deux langues. L'analyse de ce corpus, les résultats obtenus et leur interprétation feront l'objet du chapitre six.

Dans le septième chapitre, nous allons présenter une étude sur les méthodes de représentation structurée de discours que nous appliquerons ensuite à nos corpus dialogiques.

En conclusion, nous mettrons en évidence les résultats importants de ce travail.

II. Vers la Communication Homme-Machine (CHM)

Pour concevoir et développer des systèmes capables de dialoguer avec un interlocuteur humain de manière efficace et adaptée, il est nécessaire au préalable d'observer les phénomènes mis en jeu au cours d'une communication entre des humains.

Le moyen le plus connu chez l'homme pour interagir est le langage.

Le langage est une notion vaste, qui englobe tout ce qui communique ou exprime quelque chose. C'est donc un ensemble de signes permettant de s'exprimer, de communiquer entre plusieurs entités. Le langage est une aptitude innée à communiquer propre à l'être humain. Les communautés se regroupent autour d'une langue.

La langue représente un système fixé de signes différenciés propre à un groupe permettant l'expression et la communication d'informations. C'est donc un code constitué par un ensemble de règles communes à cette même communauté.

La parole quand à elle, elle est l'utilisation individuelle du langage dans une langue déterminée, afin de dire quelque chose (généralement à quelqu'un); elle ne suppose pas essentiellement la voix.

En conclusion, nous pouvons dire que le langage chez l'homme, est la capacité d'exprimer une pensée et de communiquer au moyen d'un système de signes.

II.1 Actes de langages

Pour Austin (1962) l'usage du langage t représente des actes de discours.

Il distingue Trois types d'acte de langage/

- 1) **Acte** locutoire → : communique l'information contenue dans l'énoncé.
- 2) **Acte** illocutoire → est fait avec une certaine force (force illocutoire) qui porte une intention.
- 3) **Acte** perlocutoire → provoque certains effets sur l'auditeur . On parle dans le dessein de provoquer ces effets

Les actes de langage sont composés d'une force illocutoire et d'un contenu propositionnel : F(P)

On voit donc que : (J. L. Austin, 1962 How to do Things with Words)

- **pour la pragmatique** : On se réfère uniquement à celui qui parle, ou, en terme plus général, aux usagers du **langage**.
- **pour la sémantique** : Nous faisons abstraction des usagers du langage et nous analysons seulement les expressions et leurs **significations**.
- **pour la syntaxe** : Nous faisons abstraction des significations pour analyser uniquement les relations **entre expressions**,

La totalité de **la science du langage**, se composant des trois parties mentionnées, forme la **sémiotique**

Il apparaît alors une conception pragmatique du **langage**

II.2 La pragmatique?

La pragmatique s'occupe non pas des " aspects codiques du langage" (affaire de la linguistique), mais uniquement des "processus d'interprétation qui viennent se superposer au code pour livrer une interprétation complète des phrases"(Reboul et Moeschler, 1998)

La pragmatique pour objet d'étude l'énoncé vs la phrase.

La phrase est l'unité maximale de la linguistique, composée de morphèmes (lexicaux et grammaticaux), eux-mêmes composés de phonèmes.

La phrase a une ou plusieurs significations.

L'énoncé est l'occurrence d'une phrase par un locuteur dans une situation particulière.

L'énoncé a un sens dans un contexte. La pragmatique intervient pour

Rendre compte de l'usage du langage dans la communication.

- Expliquer comment les inférences fonctionnent dans l'interprétation des énoncés.
- Expliquer pourquoi la communication non-littérale est préférée à la communication littérale.

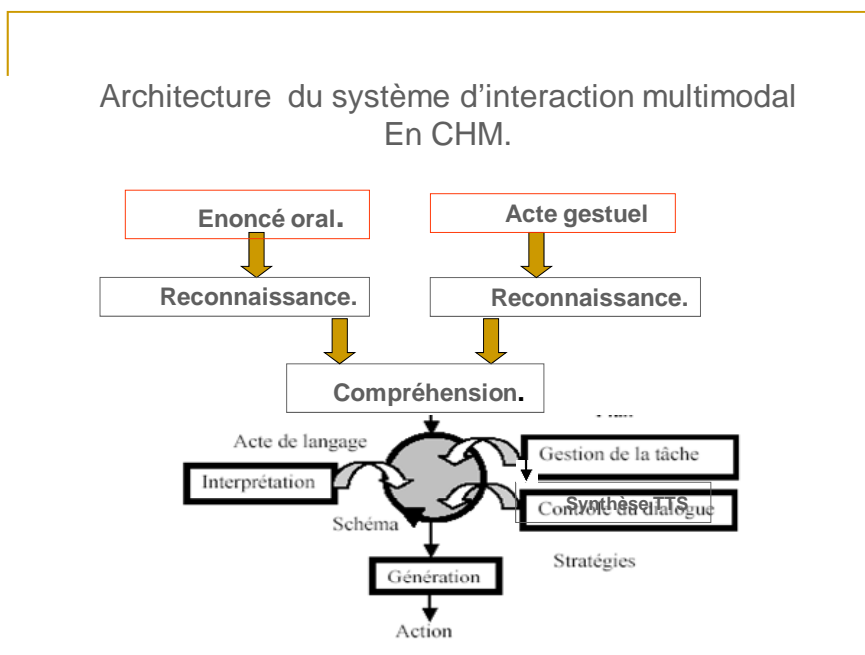
- Expliquer la sous-détermination linguistique de l'interprétation des énoncés

La pragmatique est complémentaire à la linguistique et traite les **phénomènes** tels que :

- Désambiguïsation : Choix d'une interprétation relativement au contexte
- Enrichissement : Attribution d'un référent aux Expressions Référentielles
- Attribution d'une Force Illocutionnaire aux énoncés
- Attribution d'une Attitude Propositionnelle au locuteur
- Inférences
- Implications (conclusions implicites)
- Formation des hypothèses contextuelles (prémises implicites)

En conclusion ce sont tous ces aspects que nous voulons modéliser pour construire des interfaces homme machine de haute qualité.

III. Architecture du système d'interaction multimodal En CHM.



Le système CHM est une chaîne complète depuis le module de reconnaissance au module de synthèse de la parole.

Le module RAP traduit l'énoncé sonore en une chaîne orthographique de mots. En parallèle le module de reconnaissance des gestes construit une structure morphologique de l'action gestuelle. Cette chaîne orthographique et cette structure gestuelle sont pris en charge par le module de compréhension qui fournit un schéma Sémantique représentant le sens littéral de l'énoncé multimodal.

Ce schéma sémantique s'enrichit au fur et à mesure des apports des autres modules pour devenir un acte de langage muni d'un contenu propositionnel, grâce aux apports du module d'interprétation et du module de gestion de la tâche.

Cet acte de langage est envoyé au contrôleur de dialogue qui détermine alors le but à atteindre relativement à la tâche en cours, la stratégie de dialogue à suivre et le type de réponse à faire.

Il charge le module de génération de construire la forme linguistique finale transformée en chaîne vocale par le synthétiseur TTS pour la composante communicative et la forme actionnelle pour la composante activée.

III.1 Le système de reconnaissance automatique de la parole RAP

Il permet de traduire les signaux sonores en entrée par une chaîne orthographique en sortie.

Les systèmes RAP modernes reposent sur une architecture séquentielle et modulaire dans laquelle l'onde acoustique de parole est mesurée et analysée afin d'en extraire son contenu linguistique sous la forme d'une séquence de mots.

Notre choix s'est porté sur le Dragon Naturally Speaking (DNS), qui est un système de dictée continue multifonction doté d'un vocabulaire très étendu.

Dragon Naturally Speaking possède plusieurs sources d'informations :

- Un modèle acoustique
- Un vocabulaire
- Un modèle linguistique

III.2 Le Module de Reconnaissance des Gestes (RAG)

Le module de reconnaissance des gestes construit une structure morphologique de l'action gestuelle.

III.3 Module de compréhension

A partir de la chaîne textuelle produite par le module RAP et la structure morphologique produite par le module RAG, le module de compréhension fournit un schéma sémantique à sa sortie il réalise ainsi la fusion de deux événements. Il doit donc être robuste et avoir une faible sensibilité aux erreurs venant du module RAP. Sa qualité a un impact direct sur le système. Des approches différentes ont été proposées pour ce module selon les systèmes. En général, nous utilisons une méthode qui effectue les deux analyses successivement l'analyse syntaxique et l'analyse sémantique.

III.4 Le gestionnaire de la tâche

Le contrôleur de la tâche prend en entrée le but du dialogue à résoudre et développe un plan pour l'atteindre.

III.5 La génération d'énoncé

Le module de génération prend en entrée les consignes (un acte de la machine, but de dialogue, stratégie de dialogue) du module de contrôleur de dialogue et fournit en sortie une suite d'instructions exécutables et/ou produit une chaîne orthographique textuelle.

III.6 Le synthétiseur de la parole

La tâche du synthétiseur de parole est de convertir une chaîne orthographique en un signal sonore (TTS : tex to speech). Le temps de synthèse TTS doit être suffisamment rapide pour ne pas augmenter le temps de l'interaction entre l'homme et la machine.

IV. Un système de gestion d'évènements multimodaux.

4.1 Introduction

Il est bien admis maintenant que la multimodalité peut être vue sous deux aspects : l'un est relatif aux aptitudes et à la volonté de l'utilisateur, dans ce cas, il est intéressant de voir ce problème sous l'angle de la psychologie, de l'ergonomie cognitive et des rapports humains naturels...etc. L'autre est relatif au système, comment la machine doit se comporter vis-à-vis de l'utilisateur ? Et ceci indépendamment de la méthode adoptée.

Pour bien comprendre ces phénomènes, il faudra passer par l'expérimentation, c'est pourquoi, nous avons voulu monter un système qui va mettre en situation de communication l'utilisateur et la machine.

4.2 Description de notre application.

Cette application consiste à faire remplir un document scolaire par la machine à l'aide d'un dialogue multimodal. Les modalités que nous avons prises sont la parole et le geste de pointage à l'aide de la souris. Nous avons fait de sorte que chaque modalité peut être utilisée seule ou combinée avec l'autre. L'un des documents que l'on peut remplir est le certificat de scolarité de l'étudiant.

Info	Tâches	Nouveau	Exemple	Aide	Quitter
Entrer Remplacer Effacer					
NOMS		NIVEAU		CERTIFICAT DE SCOLARITE	
BENIF ZINEB TOURI MOURAD YANES REDA		3ème ANNEE 4ème ANNEE 5ème ANNEE		UNIVERSITE DES SCIENCES ET DE LA TECHNOLOGIE HOUARI BOUMEDIENE U.S.T.H.B	
OPTION		OK		Nom: Prénom: Option:	
COMMUNICATION CONTR OLE INSTRUMENTATION					
Bienvenue au service de la scolarité! Que désirez-vous?					

Sur ce document, nous nous sommes limités à trois champs à remplir et les éléments servant au remplissage sont affichés à l'écran sur trois listes différentes. On peut réaliser les trois tâches : entrer, remplacer, effacer. Nous avons deux fenêtres de dialogue. La première permet à l'utilisateur d'interroger la machine ou de mener une négociation avec celle-ci. Voici un exemple de dialogue que l'on peut avoir avec la machine.

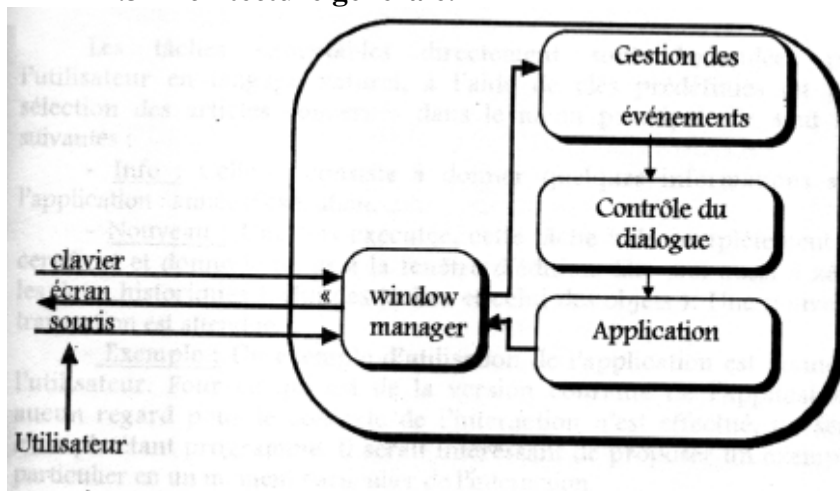
M : que voulez vous faire ?

U: entrer un nom

M: oui, le quel ?

U: Réda
M: oui, et ensuite que voulez-vous faire ?
U: entrer Zineb
M: Il y a déjà un nom sur le certificat, voulez-vous le modifier ?
U: non, le niveau de l'étudiant est celui de la quatrième année
M: oui, et maintenant ?...etc
U: je veux remplir un autre certificat

4.3 Architecture générale.



Le système a une structure permettant, grâce à Windows manager, de simuler le parallélisme des événements qui arrivent au gestionnaire, en effet c'est cette possibilité qui a fait qu'on peut combiner les deux événements parole et le clic de la souris.

4.2 Modèle de l'application

Nous allons examiner les objets et les tâches relatifs à notre application. Les objets ont le sens que Windows leur donne, ce sont des fenêtres. Nous les avons structurées sous forme d'un arbre, les fenêtres objets sont organisées en fenêtres mère et fenêtres filles. C'est la structure de la programmation orientée objet qui nous permet de représenter nos objets.

Par exemple :

Objet : nom

Attr = nom propre

Valeur = {Bouzid réda, Rami mourad, ...kaci ali}

Objet : niveau

Attr = niveau

Valeur = {1,2,3,4,5}

Objet : option

Attr = option

Valeur = {contrôle, instrumentation, communication}

Cette technique est aussi utilisée pour définir les tâches que l'on peut réaliser dans notre application, c'est-à-dire les tâches donnant lieu à des actions exécutables directement et celles qui passent par le contrôleur de dialogue.

Les tâches exécutables directement sont demandées par l'utilisateur soit à l'aide de la parole, soit à l'aide de clefs prédéfinies ou par sélection des icônes présentées dans le menu principal

Les fonctionnalités principales sont :

- Info : cette fonction retourne des informations sur l'application.
- Nouveau : on demande de vider le certificat pour remplir un nouveau.
- Exemple : on demande un exemplaire de l'utilisation de notre application.
- Aide : c'est par cette rubrique que l'on peut apprendre l'usage de notre application.
- Quitter : on sort de l'application.

Les tâches "entrer", "remplacer" et "effacer" passent par le contrôleur de dialogue.

On peut donc considérer le niveau linguistique de notre application sous forme de deux modules :

- l'analyseur du langage naturel dont le rôle est de produire une représentation interne, pour la machine, du sens des phrases de l'utilisateur.
- le traducteur dont le rôle est de transformer la représentation interne en une action ou une suite d'actions pouvant être exécutée par le système. Ce système utilise des connaissances lexicales (nom, verbe, ...), morphologiques (genre, nombre, temps, mode ...) et sémantiques (sens, concepts, lien entre concepts ...). Nous avons besoin donc d'un dictionnaire qu'il faudra structurer, pour cela plusieurs possibilités, la première, la plus simple serait de représenter toutes les formes fléchies des mots. L'inconvénient majeur est la taille volumineuse du lexique. Mais elle a un grand avantage, elle ne nécessite pas de traitement morphologique au niveau du mot. En effet la reconnaissance d'un mot ne passe pas par la reconnaissance des désinences qui sont les marques de fin de mot permettant de définir les genres et nombres des noms et des adjectifs, et les temps, le mode et personne des verbes. Cette méthode permet à l'analyseur de se contenter de rechercher le mot qui apparaît dans le texte introduit en entrée. Une autre méthode serait de découper le mot, en "sous mot" racine, préfixe et suffixe. Mais comme on a simplifié volontairement notre application, nous avons utilisé la première méthode qui consiste à prévoir dans le lexique tous les mots fonctionnels et surtout pouvoir prévenir toutes les formes qu'ils pourraient prendre dans les l'énoncé de l'utilisateur.

L'espace "objets-tâches" de cette application contient les mots fonctionnels suivants :

{Information, aide, exemple, nouveau, quitter, entrer, remplacer, effacer, nom, option, niveau, 1, 2, 3, 4, 5, contrôle, instrumentation, liste des noms et prénoms des étudiants...etc.}

Il faudra représenter maintenant tous les synonymes, il est donc nécessaire d'établir des relations entre les mots utilisés. Nous avons deux possibilités :

- On établit une relation symétrique entre les deux mots, en indiquant que le premier a pour synonyme le deuxième ou vice versa.
- on choisit l'un des deux comme référence et on renvoie l'autre à cette référence, cette relation n'est plus symétrique.

La première méthode est difficile à mettre en œuvre, nous avons préféré la seconde, car outre le gain de place, elle permet un accès rapide au synonyme. Nous avons pris les mots clefs de l'application comme référence pour définir leurs synonymes. Il reste à définir maintenant les catégories syntaxiques du langage utilisé en fonction des catégories sémantiques de l'application.

4.2.1 Catégorisation des mots du dictionnaire

Le domaine de notre application étant relativement limité, il fournit un cadre où les relations entre objets deviennent faciles à établir car elles sont connues, c'est pourquoi nous allons construire des catégories syntaxiques orientées vers le domaine étudié plutôt que de passer par l'usage d'une grammaire.

Pour simplifier cette catégorisation on doit restreindre le champ couvert par le vocabulaire employé et le nombre de constructions syntaxiques permises.

Dans la version actuelle, nous avons réduit au maximum les catégories sémantiques en :

-catégorie <ensemble>

Elle désigne l'ensemble des classes de mots qui font référence aux entités existantes

-catégorie <attribut>

C'est l'ensemble des classes de mots qui font références aux rubriques existantes

Exemples : "nom", "sexe", "numéro"...

Dans cette application, nous avons les classes sémantiques suivantes :

-la classe sémantique 'nom'

-la classe sémantique 'option'

-la classe sémantique 'niveau'

-catégorie <valeur>

C'est l'ensemble des classes de valeurs que peuvent prendre les attributs

Exemples :

"nom" admet pour valeur une chaîne de caractères,

"niveau" admet pour valeur une valeur numérique ({1, 2, 3, 4,5})

"option" admet pour valeur une chaîne de caractères ({contrôle, communication, instrumentation},...).

La catégorie syntaxique est donc subdivisée en trois catégories sémantiques : ensemble, attribut et valeur.

Notre application étant relativement limitée, notre lexique est constitué à peu près de 300 mots environ dont cinquante sont significatifs : ce sont les mots clefs, les concepts ou les références.

Le lexique a la structure d'un arbre.

Les mots significatifs sont :

E(1) : action _ exécutable = {information, aide, exemple, nouveau, quitter}

E(2) : action _ non _ exécutable directement = {entrer, remplacer, effacer}

E(3) : attribut = {nom, option, niveau};

E(4) : valeur _ niveau = {1, 2, 3, 4,5};

valeur _ option = {contrôle, communication, instrumentation}

valeur _ option = {Réda, Zineb, Amar, Rachida}

E(5) : ensemble = {étudiant,...}

E(6) : det = {le, la, tout,...}

Traitement des aspects linguistiques et gestuels de l'interaction

Lorsqu'on rentre une chaîne de caractères, l'analyse détermine les unités linguistiques et doit arriver aux unités syntaxico-sémantiques obtenues par composition morphologique des unités linguistiques. Pour identifier ces unités, il faudra donc mettre en œuvre un analyseur morpho_lexical et un analyseur syntaxico_sémantique. Le premier permettra d'obtenir toutes les définitions possibles de chaque mot avec ses différents traits. Le deuxième construira grâce à une analyse dirigée par la syntaxe, une représentation interne du sens des phrases traitées lorsque celles-ci sont correctes.

Analyseur morpho_lexical

Le clavier ou le système de reconnaissance de la parole permet d'entrer un énoncé sous forme d'une chaîne de caractères. Celle-ci reste visible sur la fenêtre d'édition jusqu'à ce l'on appuie sur la touche « retourne » alors un macro événement parole est créé. C'est à ce moment là, que celui-ci appelle le module qui s'occupe de l'analyse morpho_lexical qui recherche toutes

les unités linguistiques qu'il range dans un vecteur. Comme on travaille en mots clef, nous faisons un filtrage pour ne retenir que ceux du dictionnaire. On obtient alors des mots contenus dans le lexique de notre application, qui renferme tous les éléments que nous avons jugé pertinents au bon déroulement de la communication. Le vecteur des mots de la phrase est finalement obtenu.

Analyseur syntaxico_sémantique.

Il s'agit d'utiliser la représentation interne des connaissances de notre application pour l'analyse syntaxique. Nous avons choisi la grammaire des cas de FILMORE. Après avoir filtré les éléments non pertinents de l'énoncé, une analyse approfondie des parties essentielles est menée car si la phrase est longue nous avons opté de faire une composition de la signification de la phrase à l'aide de celles des mots clefs, l'analyse repose donc sur ces mots.

Dans l'exemple:

- Entrer s'il vous plait le nom Réda sur le certificat
- Entrer Réda

Les deux phrases sont équivalentes dans notre analyse puisque seuls les mots clefs « Entrer » et « Réda » sont examinés. L'analyse morpho_lexical donne un vecteur de mots clefs puis un tableau relatif au clavier est rempli par l'analyseur syntaxique, ce tableau contient une traduction des mots en la représentation interne des objets et des tâches du système.

Si l'on examine l'exemple suivant :

- " Entrer le nom Réda"

Anal_tâche = entrer

Anal_tâche.objet = Réda

Anal_obj.attr = nom

Anal_obj.valeur = Yanés Réda

Destination = position équivalente sur le certificat

Dans ce cas on ne manipule pas la position des objets de l'espace « certificat »

- « Remplacer le nom Réda par Mourad »

Anal_tâche = remplacer

Anal_obj1.attr = nom

Anal_obj2.valeur = Touri mourad

Lieu = position correspondante sur le certificat

- « effacer tout »

Anal_tâche = effacer

Lieu = certificat

Anal_tâche.det = tout

L'analyseur est donc capable de

1) retrouver les mots clefs dans la commande et traiter les synonymes (exemple : mettez = entrez),

2) instancier à l'aide de ces mots les bons attributs dans le tableau d'analyse.

L'analyse syntaxique basée sur la grammaire des cas permet de se focaliser sur le prédicat pour attribuer un sens à la phrase suivant les cas qui entoure ce dernier. Il s'agit en fait de traiter les « tâches » par notre application, on aurait pu interpréter symboliquement la phrase et on aura alors une vision abstraite mais il faudra revenir ensuite au contexte étudié. Dans ce cas, on s'approche mieux des aspects linguistiques mais en contrôle de processus, les applications ne demandent pas tant. On cherche plutôt un gain de temps, c'est possible car comme dit FIMORE,

on pense directement « application » et encore plus directement « tâches permises par l'application ».

L'analyse effectuée par les modules morpho lexical et syntaxico_sémantique ont permis de remplir le tableau d'analyse réservé au clavier.

4.4. Traitement des aspects gestuels

Dans cette étude, on utilise la souris pour agir sur l'espace "document", en effet par la désignation du clic souris on peut accéder à un menu, sélectionner un objet de l'application mais aussi de déplacer un objet d'une position vers une autre, comme par exemple si on clique sur un élément d'une liste, suivi d'un autre clic sur une position dans le document, on déplace l'élément vers la position sur le document. Nous utilisons la même démarche pour analyser les différents gestes, et traduire en termes de représentations internes des objets et des tâches de l'application.

L'analyse effectuée sur les événements gestuels ont permis de remplir le tableau d'analyse réservé à la souris.

Prenons quelques exemples d'analyse des énoncés concernant les composantes gestuelles :

- 1) L'utilisateur clique sur Réda
Anal_obj.attr = nom
Anal_obj.valeur = Réda
- 2) L'utilisateur clique sur Réda puis sur certificat
Anal_tâche = entrer
Anal_obj.attr = nom
Anal_obj.valeur = Réda
Lieu = position correspondante sur le certificat

Nous pouvons dire que l'analyseur est capable de reconnaître l'objet pointé par la souris, la tâche désignée et d'instancier les bons attributs dans la grille d'analyse correspondante.

Nous venons de voir que notre application s'appuie sur deux éléments essentiels des deux modalités utilisées : 1) les représentations internes de sens pour la parole et 2) pour le geste de désignation.

Pour mener à terme la tâche demandée par l'utilisateur, la machine doit communiquer et interagir avec celui-ci, c'est pourquoi, il est nécessaire de mettre au point un système dialogique.

4.5 Le modèle de dialogue.

De manière globale on peut dire qu'un dialogue est une conversation entre plusieurs partenaires sur un sujet en vue d'aboutir à un accord ; c'est aussi le contenu de la conversation : les propos échangés. Le dialogue se manifeste par un ensemble d'expressions dont l'émission est contrôlée de manière répartie par un ensemble d'interlocuteurs.

Dans notre application, le dialogue est dirigé par le "but", il donne l'initiative à l'utilisateur pour fixer le cadre de la tâche élémentaire mais c'est la machine qui dirige le dialogue lors des négociations utiles à la définition complète de la tâche. Il est donc nécessaire de mettre au point plusieurs modèles, modèle de l'application, de l'utilisateur, de la tâche et du dialogue..., et une méthode pour gérer les événements multimodaux et pouvoir intégrer les informations vocales et gestuelles dans une représentation commune.

Pour réaliser notre modèle, nous nous sommes inspirés du modèle à deux axes de Luzati et le modèle script de dialogue. Le premier permet la gestion du canal de communication, c'est-à-dire qu'il vise à éviter les problèmes d'intercompréhension et d'incommunicabilité en insistant sur le bon fonctionnement de l'interaction plutôt que sur le raffinement. Le deuxième permet la

représentation en orienté-objet, des objets de l'application, il se présente sous forme algorithmique.

Par exemple, voici un dialogue avec des tours de parole scrupuleusement respectés

M1 : que voulez vous faire ?	Question Principale
U1: modifier le nom	Réponse Principale
M2: Quel nom ?	QS
U2 : Zineb	RS
M3: d'accord, et maintenant ?	QP

On peut représenter ce dialogue par :

M1→U1→M2→U2→M3...ou par QP (RP (QS RS)) QP

Examinons cet autre exemple :

M1 : que voulez vous faire ?	Question Principale	
U1 : entrer	Réponse Principale	
M2: Que voulez vous entrer ?		QS
U2 : L'option de l'étudiant		RS
M3: Quelle option voulez vous entrer ?	QI	
U3: Contrôle	RI	
M4: d'accord, et maintenant ?		
	QP	

On peut représenter ce dialogue de la manière suivante

M1→U1→M2→U2

M3

U3→M4

L'autre représentation est : QP (RP (QS RS (QI RI)) QP

C'est la représentation dynamique de Luzati selon les deux axes. Cette première réalisation a permis une analyse des deux modalités de communication : parole et geste de pointage souris dans le cadre d'un traitement d'informations administratives.

Dans la réalisation qui suit, nous nous sommes intéressés plus à un dialogue avec une application graphique.

V. Les Dialogue homme-machine (DHM)

1. Introduction

La réalisation des systèmes de dialogue homme-machine (DHM) peut se faire selon cinq approches: dialogue fondé sur les attentes du locuteur, la planification, la structuration, la négociation et l'interaction. La première approche est fondée sur les attentes qui conditionnent

l'analyse de la situation du dialogue, La deuxième approche est basée sur la reconnaissance et la construction de buts et de plans : c'est la planification intentionnelle. La troisième approche est issue d'un courant structuraliste qui s'est donné pour objectif de dégager et d'expliquer la structure du discours. La quatrième approche tente de modéliser la façon dont les protagonistes d'une négociation négocient leurs enjeux conversationnels afin d'aboutir à un accord. La cinquième et dernière approche repose sur la modélisation de principes d'interprétation et de production des actions communicatives. Ces présentations, qui ne prétendent pas être exhaustives, montre les travaux marquants sur le dialogue homme-machine et rejoint, en certains cas, les travaux sur le dialogue humain issus de la linguistique et des sciences cognitives.

2 Le dialogue homme-machine, état de l'art.

Les premiers systèmes du DHM ont centré leur attention sur une simple paire question-réponse [Villaseñor-Pineda, 1999]. Dans ces systèmes, il n'y pas de véritable dialogue. L'interaction a été réduite à l'élaboration d'une réponse à une question posée. Ainsi, l'interaction a été fortement contrainte et le fonctionnement correct du système dépendait de la pertinence des informations présentées par l'utilisateur au cours d'un tel dialogue.

D'après Villaseñor-Pineda, le premier système à aborder des questions intéressantes pour le dialogue a été le système SHRDLU en 1971. Le système était capable de mener un dialogue dans un domaine d'application réduit facilement maniable. Le système SHRDLU simulait un bras mécanique, dans un monde de blocs, que l'utilisateur pouvait commander pour réaliser des actions élémentaires (par exemple : saisir un bloc, l'empiler sur autre, etc.) et les résultats de ces actions étaient visualisés dans un écran. Nous présentons sur la figure IV.I. Un exemple d'un dialogue avec le système SHRDLU.

U: *Pick up a big red block*
 S: *ok*
 U: *Grasp the pyramid*
 S: *I don't understand which pyramid you mean*
 ...

Figure IV. 1 Exemple d'un dialogue avec SHRDLU [Villaseñor-Pineda, 1999].

Un autre système important dans les premiers pas du DHM était le système GUS en 1977. GUS est un système de renseignements de vols aériens. Le tableau IV.1 montre un dialogue avec GUS. Le dialogue dans ce cas était destiné à remplir un schéma sous-jacent, qui est à l'origine des interventions de GUS et c'est la base de l'interprétation de celles de **l'utilisateur**.

Bien que GUS soit capable de réaliser un dialogue simple, l'utilisation de schémas impose un cadre rigide qui empêche de gérer des phénomènes propres du dialogue comme le malentendu ou l'incompréhension

<i>GUS</i>	Bonjour, je suis GUS, je peux vous aider à effectuer un voyage simple en avion. Où voulez-vous aller ?
<i>Client</i>	Je veux aller à San Diego, le 28 mai.
<i>GUS</i>	A quelle heure voulez-vous partir ?

<i>Client</i>	Je dois être à San Diego avant 10 heures.
<i>GUS</i>	Le vol 102 de P.S.A. qui arrive à 9h15 vous convient-il ?
<i>Client</i>	Quel est le vol suivant ?
<i>GUS</i>	Le vol 310 d'Air Californie, qui part à 8h30 et arrive à 10h
<i>Client</i>	Je prendrai le premier.
<i>GUS</i>	A quel nom est-ce que je fais la réservation ?
<i>Client</i>	Dan Bobrow
<i>GUS</i>	J'ai confirmé le vol suivant : vol 102 de P.S.A. mercredi 28 mai de San José à San Diego qui part à 7h30 et arrive à 9h15 Quel jour voulez-vous revenir ?
<i>Client</i>	vendredi soir
<i>GUS</i>	Le vol qui part à 19h45 vous conviendrait-il ?
<i>Client</i>	Parfait
<i>GUS</i>	J'ai confirmé le vol suivant : vol 307 de P.S.A. vendredi 30 mai de San Diego à San José qui part à 19h45 et arrive à 21h30 Au revoir

Tableau IV. 1. Exemple de dialogue avec GUS [Villaseñor-Pineda, 1999].

Ces premiers systèmes ont mis en évidence la nécessité de déterminer le rôle de l'information extra linguistique dans le traitement du dialogue. Nous présentons dans ce qui suit les différentes approches sur lesquelles sont basés les DHM.

Les approches fondées sur les attentes du locuteur

Dans le cadre du dialogue Lemeunier [Lemeunier, 2000 a], Lehuen [Lehuen, 1997] et Vilnat (d'après Fouquet [Fouquet,2002]), proposent d'introduire la notion d'*attentes* : Pour Lehuen, ces attentes sont "*ce qui est attendu par le système au regard de l'état courant de la tâche*". Il les définit donc en terme de "l'objet du domaine d'application et d'actes de dialogue orientés par la tâche" ; pour Vilnat, ce sont "*ce que l'utilisateur est capable 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*" ; Pour Lemeunier, ce sont "*des entités pragmatiques manipulées par le système qui correspondent à ce que l'utilisateur est censé dire et ce à quoi le système est donc censé s'attendre*". Ces attentes sont donc très liées à la tâche d'une part et considérées du point de vue du système d'autre part.

Les approches fondées sur la planification

Les modèles de dialogue fondés sur la planification partent de l'hypothèse que les personnes engagées dans un dialogue ont des comportements rationnels, qui les mènent à construire et à exécuter des plans pour atteindre des buts. Ces plans sont indifféremment constitués d'actions applicatives et d'actions communicatives. Dans ces modèles, la production d'un énoncé par un locuteur correspond alors à la réalisation d'un sous-but communicationnel planifié

[Lehuen,1997], Par exemple le modèle de Litman [Latapy & al, 2003] contrôle des interactions entre des plans répartis en plan du domaine, actes de langage et métaplans du discours.

Les approches fondées sur la structuration

Ces modèles ont été conçus pour étudier et comprendre les mécanismes du dialogue à travers sa structure, c'est-à-dire qu'ils décrivent le dialogue en définissant ses constituants et la structure qui spécifie les liens de composition, de succession ou de subordination entre eux [Villaseñor-Pineda,1999]. Le modèle genevois de Roulet en 1985 est le plus représentatif parmi les travaux de cette approche. Ce modèle permet d'identifier a posteriori les liens hiérarchiques et fonctionnels entre les constituants du discours. L'analyse hiérarchique du modèle genevois repose sur plusieurs niveaux de constituants : les incursions sont les plus grandes unités dans un dialogue, les échanges sont les plus petites unités dialogiques, les interventions sont les constituants uniques des échanges et les actes de langage les plus petites unités monologiques. On prend comme exemple le Modèle de Roulet et Moeschler en 1985 (voir [Latapy & al,2003]) qui permet :

- ✓ L'identification des constituants du discours.
- ✓ La construction des liens hiérarchiques et fonctionnels entre eux.
- ✓ Plusieurs niveaux de constituants discursifs : incursions, échanges, interventions, actes de langage.
- ✓ Complétude interactionnelle.
- ✓ Complétude interactive.

Le modèle Genevois

Le modèle genevois est à la source de la plupart des modèles structurels [Caelen, 2002]. Il est souvent utilisé en DHM car il est d'une implémentation claire et commode. Il se prête bien à une formalisation de type grammaire de dialogue. Le modèle genevois est le produit d'une collaboration pluridisciplinaire (linguistique et intelligence artificielle). Il a pour objectif de fournir un modèle utilisable pour la conception d'un système de DHM à caractère explicatif. Ce modèle postule que les constituants du discours entretiennent entre eux des relations de dépendance hiérarchique, à l'instar des constituants de la phrase. Ces dépendances sont établies sur la base de caractéristiques sémantiques et pragmatiques, déterminées par les relations intervenant entre les constituants du dialogue.

Les approches fondées sur la négociation

La fonction première de cette catégorie de modèles n'est, ni de planifier des intentions ou des croyances, ni de structurer des interventions, mais de maîtriser des enjeux conversationnels. D'après Pasquier [Pasquier & al, 2002], une négociation est un processus de discussion engagé afin d'aboutir à un accord sur un sujet quelconque. Cette définition, relativement générale, ne permet pas de se faire une idée très précise de ce qu'est une négociation.

Selon Roulet, toute négociation a sa source dans un problème qui donne lieu à une initiative du locuteur ; cette initiative appelle une réaction de l'interlocuteur, réaction qui peut être favorable ou défavorable. Si elle est favorable, le locuteur peut clore la négociation en exprimant à son tour son accord (complétude interactionnelle).

En plaçant le concept de négociation au centre de la définition du modèle hiérarchique, Roulet présente la négociation d'un point de vue essentiellement structurel.

Les approches fondées sur l'interaction

Il est assez difficile d'expliquer l'apparition d'une intention de communication : d'expliquer à chaque instant d'une conversation, pourquoi *X* a dit ceci et pourquoi *Y* a répondu cela. Les modèles fondés sur la planification tentent de le faire, mais en s'appuyant sur un formalisme qui reste très près de la tâche [Lehuen, 1997]. Lehuen définit un principe *interactionnel* comme étant une règle locale d'analyse ou de production d'une action communicative, dont les conséquences sont observables dans des corpus de dialogues.

Conclusion

Nous venons de présenter dans ce chapitre d'une manière très schématique, une synthèse bibliographique effectuée sur les principes théoriques fondamentaux et les procédés pratiques qui régissent l'évolution des systèmes de dialogue homme-machine. Nous avons relevé les contributions ainsi que les limites de chacune de ces théories vis-à-vis des avancées dans la modélisation de dialogue en situation de conception. Cela nous permettra de référencer notre modèle par rapport aux systèmes existants.

Dans ce qui suit nous allons proposer le modèle de dialogue homme-machine finalisé pour lequel nous avons opté dans notre travail : un modèle qui est orienté pour la réalisation d'un centre de renseignement scolaire, visant à recueillir un corpus en véritable dialogue homme-machine.

Pour progresser vers notre objectif, nous avons suivi la démarche suivante :

- Nous avons développé un système initial, puis nous l'améliorons progressivement au fur et à mesure au vu des résultats que nous obtenons en utilisation réelle du système de DHM.
- A chaque analyse du corpus nous devons identifier les situations nouvelles et les intégrer au module de traitement situé dans le système, pour augmenter la base de connaissance.

V. Modélisation du dialogue

1. Introduction

Pour intégrer les représentations des connaissances nécessaires au fonctionnement de notre modèle, nous avons commencé par réaliser un système de base simple (question-réponse), puis nous avons fait une simulation de la tâche de recherche des notes assistée par ce système de dialogue. Cela nous a permis de saisir les représentations conceptuelles du sujet et d'y associer des réponses. Nous procéderons ensuite au développement de notre système en utilisant une analyse morphologique de façon à regrouper des items en champs conceptuels et à vérifier qu'ils peuvent être des déclencheurs pertinents pour la tâche. En effet, il est possible d'incrémenter la base de connaissance du système de dialogue en lui ajoutant de nouvelles possibilités de compréhension contenant les formes nouvelles nécessaires au dialogue.

2. Modèle de la mémoire

Notre système de dialogue homme machine peut être considéré comme un système multi-agent, un agent gérant l'application en échangeant des informations orientant le locuteur vers la

satisfaction du but et un agent gérant l'activité langagière (l'analyse et la production des réponses). Il est constitué des connaissances statiques et des connaissances dynamiques :

-Les connaissances statiques sont des connaissances linguistiques utilisées lors de l'interprétation et de la génération. Il s'agit d'un ensemble de connaissances qui n'augmentent pas au cours de dialogue. Ces connaissances regroupent un ensemble de dérivées des transcriptions phonétiques organisées dans une mémoire.

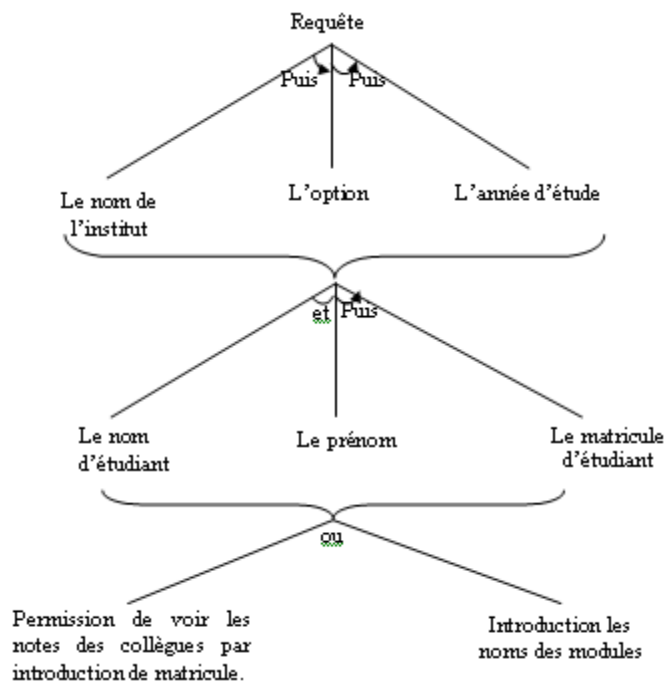
-Les connaissances dynamiques sont des connaissances portant sur l'état de la tâche et l'historique de dialogue. Ces connaissances évoluent au cours de la progression de dialogue, c'est un ensemble de dictionnaires de formes (morphèmes¹ variables construites au cours de dialogue). L'ensemble de ces connaissances permet de reconnaître l'état de la tâche par rapport à son état finale qui est la satisfaction du but.

3. Système Recherche de l'information

Notre domaine d'étude est l'analyse puis la reconnaissance des intentions de l'utilisateur. Nous allons analyser des énoncés réels en dialogue homme-machine (DHM) pour la reconnaissance des intentions qui dépend des éléments restant en attentes. Pour cela nous avons conçu une base de données qui contient l'état scolaire et l'identité des étudiants avec leurs notes d'examens. Le but de l'utilisateur est de chercher ses notes sans aucune idée ou informations préalables sur ces dernières. En dialoguant en langue naturelle, La machine tente de comprendre ses intentions et d'y répondre correctement. Après l'identification complète de l'utilisateur (dans notre travail c'est l'étudiant), la machine lui fournit les informations nécessaires en créant des commentaires sur l'information retrouvé permettant d'adapter la réponse avec l'information retrouvé.

3.1. Structure hiérarchique du système

Pour satisfaire le but, l'utilisateur doit introduire des informations concernant les connaissances dynamiques qui sont conçus et construisent comme un ensemble structuré d'arborescents élémentaires constituant les attentes. Ces attentes sont arrangées sous forme d'une hiérarchie (fig. 1) selon la priorité d'apparition.



¹ Plus petit segment portant un sens. On parle des morphèmes libres

Figure I. *Organisation des attentes.*

D'après la hiérarchie représentée sur la figure I, les attentes sont reliées à des fils par des conjonctions. Pour atteindre le but par le locuteur, il faut analyser tous les fils de la hiérarchie. L'attente la plus préférée est l'introduction de la requête c.à.d l'utilisateur vient de présenter son but; puis l'introduction des informations concernant l'état scolaire de l'utilisateur, ensuite l'introduction de l'identité ; pour finalement avoir la possibilité de consulter les informations recherchées.

3.2 Les attentes

Dans notre travail, l'attente détermine l'état de la tâche par rapport à son état final qui est la satisfaction du but. Donc les attentes présentent des éléments fondamentaux d'un acte de langage.

Un énoncé de l'utilisateur peut contenir un ou plusieurs actes de langage qui dépendent de l'état de la tâche et les informations existant.

La forme générale d'un acte de langage (Jerome 2003) est donnée par J. Searle sous l'expression $F(p)$, avec $F = \{\text{Affirmer, Demander, Promettre, Exprimer, Déclarer}\}$ et p une proposition.

Searle place la communication dans la théorie des actes de langage et plus généralement dans une théorie des intentions. Pour A, énoncer une proposition résulte de l'intention de la produire; pour B, la comprendre c'est interpréter l'intention qui l'a sous-tend, dans le contexte où elle a été produite.

Dans cette théorie, la communication se situe dans la pragmatique, et la pragmatique des actes de langage s'inscrit à son tour dans une théorie du langage et une théorie de l'action selon deux perspectives : la description des actes de langage et leur régulation selon le principe d'exprimabilité, c'est-à-dire :

- a) Énoncer des mots = effectuer des actes d'énonciation,
- b) Référencer et prédiquer = effectuer des actes propositionnels (locutoires),
- c) Affirmer, ordonner, promettre, vouloir etc. = effectuer des actes illocutoires,
- d) Effectuer des actes perlocutoires = agir sur son interlocuteur.

En (c) on peut représenter tout acte de langage par une fonction $F(p)$ (Caelen 1997).

p = contenu propositionnel et F =force illocutoire, (ex. : " je veux voir mes notes", F marquée par "je veux voir" : force vouloir et p marqué par l'information désiré "la note"). A ce niveau Searle ne prend pas en compte le rôle du destinataire ni celui du destinataire qu'il ne fait apparaître qu'en (d).

Searle propose pour l'interprétation de ce type d'acte une stratégie inférentielle (Caelen 1997), qui examine toutes les conditions de réalisation de l'acte (situation, monde, arrière-plan, etc.), le pourquoi, les intentions du demandeur, le but poursuivi, etc...

4. Le modèle coopératif

En se référant à une base de donnée, notre système permet de satisfaire le locuteur en langage naturel. Son but est de guider une conversation avec l'utilisateur humain afin de satisfaire un but personnel en gérant un dialogue homme-machine finalisé c.à.d orienté vers la réalisation d'une tâche. Il permet d'accéder à une base de connaissance en gérant une conversation c.à.d. une séquence indéterminée d'échanges sans interruption, ce qui revient à être capable de s'adapter à son interlocuteur. Notre but est de proposer un système répondant aux requêtes des utilisateurs, et aidant ces derniers pour la recherche de ses notes des examens. Au début de chaque dialogue, notre système doit saluer l'utilisateur (salut, bonjour), et lui laisser le temps et la liberté d'exprimer de. En laissant à la machine le travail d'interprétation des énoncés, et à un niveau supérieur de reconnaissance des intentions, la charge cognitive des utilisateurs est diminuée. C'est ici 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 (Lemeunier 2000 b). Ce modèle a des tactiques de conversation, ses connaissances augmentent au fur et à mesure avec la connaissance

du sujet et l'objectif de l'utilisateur, de sorte que le modèle peut réagir en fonction de la tâche en marche.

Pour nous, le dialogue est un processus de discussion commençant par une salutation du système vers l'utilisateur jusqu'à l'accord entre ce dernier et la machine à propos de la tâche qu'ils vont faire, puis atteindre un but commun. Donc un dialogue entre A (chercheur de l'information : l'utilisateur) et B (le producteur de la base de données : le système) peut être formalisé de deux manières : (Cherabit & al, 2005)

i- Dialogue (A, B) : demande (A, B) + informations (B, A) + annonçant (A, B) + être content (A, B).

Dans un premier temps, le dialogue entre A et B commence par une demande de la requête de A à B. Dans le deuxième temps, le dialogue se poursuit avec la proposition (de B à A) d'introduire les informations concernant l'identité de l'utilisateur. Dans un troisième temps, A fournit les informations proposés à B. Puis B évalue la pertinence de l'information présentée pour montrer les informations cherchées. Le dialogue peut alors se terminer, ou bien se poursuivre.

ii- Dialogue (A, B) : proposition (B, A) + vouloir (A, B) + information (B, A) + annonçant (B, A) + être content (A, B).

Le dialogue entre A et B commence à l'initiative de B, qui propose à A d'introduire sa requête. Ce dernier peut accepter ou refuser cette proposition. S'il accepte, B oriente A et lui propose de satisfaire sa requête en suivant le cas montré ci-dessus (i). S'il refuse, B oriente A et lui demande d'introduire sa requête, et le dialogue se poursuit. Au cours du dialogue le système évalue la bonne formation du dialogue en termes d'échanges.

5. Satisfaction du locuteur

Au cours d'un dialogue le système instancie toute la hiérarchie, construit progressivement plusieurs structures produisant des réponses en basant sur l'arborescent prédéterminé. A chaque instant le système calcule la pertinence des informations identifiées que contient l'énoncé afin d'orienter vers la satisfaction du but; et si le locuteur reprend la même phrase qu'une autre déjà énoncée, le système l'informe qu'il répète. Au cours des échanges, le système analyse les énoncés en se basant sur les attentes organisées progressivement sur la hiérarchie comportant des intentions.

A l'ouverture du dialogue si le système ne comprend pas l'énoncé du locuteur, il l'oriente vers le dialogue, on fournissant une réponse pouvant continuer le dialogue "eh, comme ça toujours vous ne dites pas bonjour?". Si le locuteur ne pose pas la requête au cours d'un dialogue et le système ne comprend pas un énoncé, ce dernier s'oriente vers la requête, où en lui associant une réponse : "pouvez vous m'indiquer quelle est votre requête?". Une fois l'utilisateur vient de présenter son but, la machine lui orienter directement à introduire des informations concernant son état scolaire (fg. II).

D'après la figure II et après que l'utilisateur a introduit les informations demandés et qui sont conformes à celles de la base de donnée, les éléments en attentes deviennent satisfait et la machine lui orienté de nouveau à introduire d'autres éléments qui restent en attente.

Le processus se poursuit jusqu'à l'accomplissement de tous les éléments demandés restant en attente.

La satisfaction du locuteur est étendue aux éléments de la décomposition de l'acte (Lemeunier 2000 a). Un acte de langage est réussi si et seulement si toutes les attentes entrant en jeu sont réussis. Ainsi si un élément représentant une attente n'est pas réussie, on ne peut pas passer à l'autre situe après cette dernière, par exemple on ne peut pas accéder au numéro de carte d'étudiant si les éléments de l'identité "nom et prénom" ne sont pas introduit et l'action reste en attente, mais on peut continuer le dialogue sans passer par ces éléments. Et quand le système ne comprend pas un autre énoncé il réoriente le locuteur pour réussir les éléments en attente.

Conclusion

La reconnaissance des intentions dans un dialogue homme-machine dépend des actes de

langage qui contient les attentes au regard de la machine, dépendant des informations fournis par l'utilisateur et qui sont conformes à celles de la base de donnée. Le système réalisé gère des dialogues en langue naturelle. En se basant sur les attentes (représentant des intentions) de la part de la machine et l'arborescent prédéterminé, le système satisfait le locuteur.

6. Exemples de dialogue, en arabe et en français.

20:07:21 M0:-salut, bonjour
 20:07:44 U1:-salut merci
 20:07:46 M1:-comment ça va ?
 20:08:29 U2:-bien et toi ça va ?
 20:08:30 M2:-Oh... ça va !
 20:09:06 U3:-bon je suis ici pour ce qui concerne les résultats de mes notes
 20:09:07 M3:-qu'est ce que vous pensez?
 20:09:23 U4:-mes notes.
 20:09:26 M4:-quels sont votre institut, votre option et l'année scolaire?
 20:11:00 U5:-médecine
 20:11:01 M5:- vous avez dis : médecine. ? Donc vous êtes en médecine introduisez l'option et l'année.
 20:11:24 U6:-pneumo
 20:11:27 M6:-bien, vous êtes en (médecine_pneumo). Quelle année ?
 20:11:47 U7:-cinquième année.
 20:11:49 M7:-donc, vous êtes en cinquième année médecine option pneumo. Pouvez-vous m'indiquer votre identité ?
 20:12:50 U8: - kamel
 20:12:53 M8:-Kamel est votre prénom, introduisez votre nom S.V.P
 20:13:27 U9:-khennic
 20:13:29 M9:-non, je ne peux pas vous aider si vous n'introduisez pas votre nom.
 20:13:57 U10:-khenniche
 20:13:58 M10:-D'accord.Vous êtes khenniche kamel, Introduisez votre matricule.
 20:14:10 U11:-1976
 20:14:13 M11:-très bien. Comme ça vous êtes bien identifié, maintenant vous pouvez voir vos notes. Introduisez les modules suivants : bactériologie humain. Médecine interne.
 20:15:10 U12:-bactériologie humain
 20:15:12 M12:-bravo. En bactériologie humain vous avez 17/20, continuez.
 20:15:53 U13:-médecine interne
 20:15:56 M13:-aucun effort. En médecine interne vous avez 2/20, attention.
 20:17:14 U14:-merci pour le service
 20:17:17 M14:-je vous en prie.
 20:17:49 U15:-merci et a bientôt
 20:17:50 M15:-au revoir, merci pour votre collaboration, si vous voulez me parler fermez cette fenêtre puis ouvrez de nouveau.

M:- سلام ومرحبا
 U:- سلام عليكم
 M:-? كيف حالك اليوم
 U:- لا بأس
 M:-... اوه ؟
 U:- ماذا اوه
 M:-? لماذا تسال
 U:- هل تستطيع أن أرى نقاطي?
 M:-? هل تستطيع ان تعرفني بنفسك ؟
 U:- توتو 5972673
 M:- لقد أدخلت لقبك فقط. ادخل اسمك من فضلك
 U:- يوسف

M:- جيد (توتو يوسف) هيا الآن ادخل رقم بطاقة الطالب -

U:- 5972673

M:- جيد. هكذا تستطيع أن ترى نقاطك. حتى لأصحابك
ادخل اسم المواد. مثلا: sec100, sep200, tec621
tec006, tec003, tp10tec505, sem310

U:- sem 310

M:- عندك 10

U:- tec 003

M:- عندك 14

U:- صحا

M:- دقق قليلا من فضلك

U:- شكرا

M:- لا شكر على واجب

U:- إلى اللقاء

M:- شكرا على الزيارة إن تريد شيء أغلق هذه النافذة ثم افتح من جديد.

VI. Interprétation des résultats

1. Introduction

Nous exposerons dans ce chapitre, l'analyse des résultats obtenus de dix dialogues homme-machine en communication réelle. Nous commencerons en premier par extraire les différents types de réponses utilisées par les usagers, ensuite nous exposerons l'analyse du corpus.

2. Étude du type de réponses des usagers dans le corpus

Nous avons étudié le type de réponses des usagers en fonction des questions posées par le système. Nous avons distingué six types de réponses dont la répartition dans le corpus est donnée dans la Figure VI.5.

Réponses correctes

Les réponses que nous appelons correctes sont celles qui répondent exactement à la question posée par le système sans ajouter aucune autre information. Ce type de réponse est le plus fréquent : il représente plus de "42%" des réponses (le nombre est calculé par rapport au nombre total des énoncés répondant à une question du système).

Réponses sur-informatives

Ce sont les réponses qui répondent à la question posée et complète cette réponse en y ajoutant de nouvelles informations non demandées. Du point de vue de la qualité des réponses, ce type de réponse est parfaitement correct mais nous avons souhaité les différencier des réponses correctes. Les réponses sur-informatives sont peu nombreuses "8%", ce qui montre que peu d'usagers prennent l'initiative de donner des informations que le système ne demande pas.

Réponses sous-informatives

Ce type de réponses fait référence aux énoncés des usagers qui répondent seulement en partie à la question du système. Elles sont assez fréquentes (près de 24%). Certaines réponses peuvent être à la fois sous-informatives et sur-informatives lorsque l'utilisateur ne répond que partiellement à la question mais donne en plus des informations non attendues. Cependant, ce phénomène est très rare, nous en avons relevé seulement un cas parmi 53 réponses sur les dix corpus. C'est pour cette raison que nous ne les avons pas distinguées et nous les avons classées comme des réponses sous-informatives.

Réponses ambiguës

Ce type de réponses regroupe les énoncés dont le sens est ambigu et pour lesquels il n'y a aucun moyen de lever l'ambiguïté. Un cas courant de réponse ambiguë est lorsqu'à la question portant sur l'introduction de nom de l'institut, l'utilisateur répond seulement par un nom d'une option et on ne peut savoir si c'est celle de tel ou tel institut, il y en a relativement souvent près de 18%.

Sans réponse

Ce sont les énoncés dits *vides*, c'est-à-dire lorsque l'utilisateur n'a rien répondu ou lorsque l'énoncé ne contient que des hésitations. Ce type de réponses est très peu fréquent, il y en a environ "2%".

Réponses hors-sujets

Ce sont les réponses non pertinentes par rapport à la tâche, par exemple lorsque l'utilisateur introduit une identité qui n'existe pas dans la base de données. Les réponses hors-sujet contiennent aussi les demandes hors périmètre. Elles représentent environ 6 % des réponses.

- R.C : Réponses correctes
- R.Su.I : Réponses sur-informatives.
- R.So.I : Réponses sous-informatives.
- R.A : Réponses ambiguës
- S.R : Sans réponse.
- R.H.S : Réponses hors-sujets.

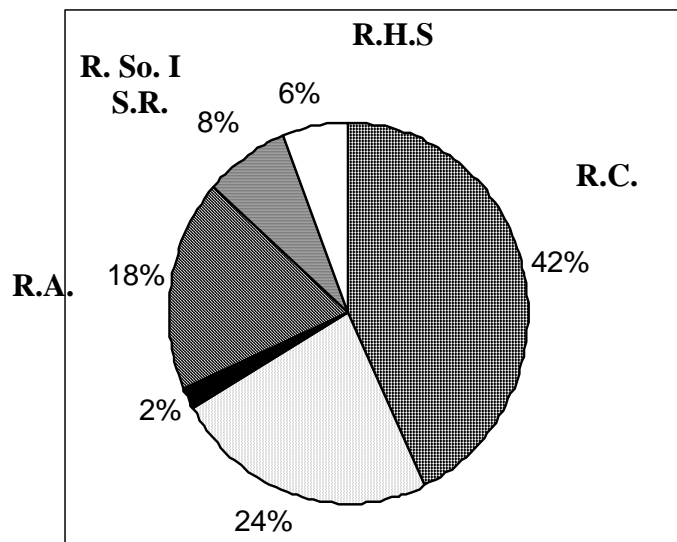


Figure1 Répartition des types de réponses des utilisateurs.

Nous observons que 50% des énoncés utilisateurs répondent à la question du système (réponses correctes et sur-informatives), ce qui représente 92% des réponses non vides (c'est-à-dire il existe une réponse à chaque question posée).

Les questions les plus répondues :

- Q1—>Pouvez-vous m'indiquer votre requête ?
- Q 2—>Quel est votre institut, votre option et votre année ?
- Q 3—>Quelle est l'année ?
- Q 4—>Venez-vous de vous présenter ?
- Q 5—>Pouvez-vous m'indiquer votre identité ?

- Série1 : réponses complètes
- Série2 : réponses incomplètes.
- 1—>Q1
- 2—>Q2
- 3—>Q3
- 4—>Q4
- 5—>Q5

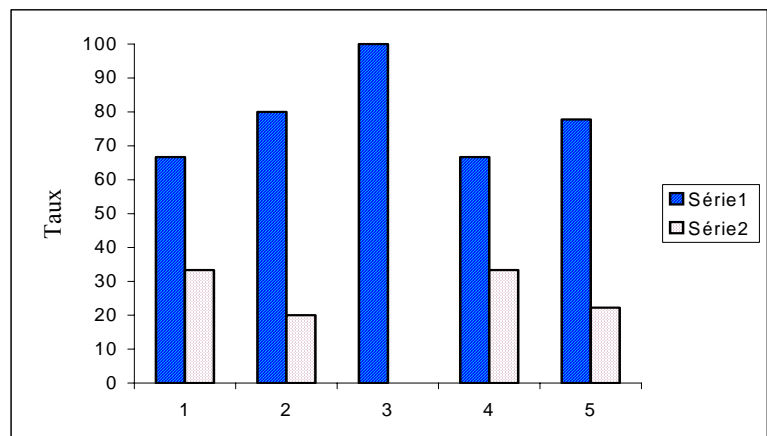


Figure2 Taux des réponses correctes et incorrectes pour les cinq questions.

3. Analyse de corpus

Nous allons présenter dans ce qui suit quelques résultats qui nous semblent intéressants. Nous présentons les critères utilisés par les usagers pour essayer d'identifier chacune des tâches de l'application. Dans l'exemple présenté dans le tableau1 ci-dessous, la demande de la requête diffère d'un usager à un autre.

Est-ce que je peux voir ma note ?
je voudrais savoir quelles sont les notes de mes dernières années.
chercher ma note
Pouvez-vous me donner mes notes ?
donner moi la note
Pourriez-vous me renseigner sur mes notes ?
je veux ma note
Je peux savoir ma note de ...
Je peux voir ma note ?
Je voudrais connaître ma note
J'ai besoin de mes notes .

Tableau1 Concept de la requête "la recherche de la note" par différents usagers

D'après le tableau1 le taux pour que l'utilisateur questionne le système concernant la possibilité ou non de voir ses notes atteint (58%), ainsi que la demande sans poser une question est (42%).

Il existe des cas de non compréhension des énoncés, où la machine tente de poser la question "*pouvez vous m'indiquer votre requête ?*", L'utilisateur répond tout simplement avec des réponses simple comme par exemple : "*voir ma note*", "*je veux ma note*", etc.

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

Nous présentons ici des éléments statistiques permettent de déterminer comment le vocabulaire employé par la machine peut influencer sur celui de l'utilisateur.

Terme	Homme (%)	Machine (%)
Note	44.4	33.3
Requête	0	33.3
Etude	11.1	9.5

Examen	22.2	4.7
Module	11.1	4.8
Cherche	11.1	14.3

Tableau 2 Quelques termes spécifiques à la tâche utilisée.

On note dans le tableau 2 que la machine énonce plus souvent le mot " requête " ; Cela implique que dans la plus part des cas, l'homme n'emploie pas un mot indiquant la requête, tandis que la machine l'intègre automatiquement au cours du dialogue pour s'orienter vers la requête.

3.2 Les verbes les plus utilisés par les usagers

Afin d'étudier la composition du corpus en terme des verbes utilisés par les usagers pour dialoguer avec la machine nous avons identifié quelques verbes les plus utilisés sur le tableau 3.

Verbes	Taux de la machine (%)	Taux de l'utilisateur (%)
Avoir	1.9	16.6
Pouvoir	4.5	8.3
Indiquer	10.9	0
Introduire	27.5	0
Être	6.4	20.7
Préciser	8.9	0
Vouloir	8.3	37.4
Parler	7.0	0
Chercher	1.9	4.1
Comprendre	4.4	0
Aider	3.2	0
Savoir	0.7	12.5

Tableau 3 Taux d'apparition des verbes les plus utilisés

Le tableau 3 montre le taux d'apparition des verbes les plus utilisés dans notre corpus. On note la présence d'un taux élevé concernant des verbes utilisés par les locuteurs et le système : "être" ; " vouloir " et " pouvoir ". Le verbe " être " est utilisé pour la personnalisation, le verbe "vouloir" apparaît lorsque l'utilisateur veut amener un message et le verbe " pouvoir " est utilisé plus souvent lorsque le système donne à l'utilisateur la permission de chercher les informations désirées.

3.3 Les substantifs utilisés par les usagers

Le tableau 4 ci-dessous présente une liste des substantifs avec le taux d'apparition.

On note d'abord que le taux d'apparition est maximal en relation avec la tâche pour le mot "note". Ce qui indique que le sujet de la tâche tourne autour de la note d'examen.

Substantif	Taux pour la machine (%)	Taux pour l'utilisateur (%)
Note	20.7	30.8
Ami	11.6	30.8
Année	18.1	15.4
module	15.6	23.1
Nom	10.3	0
Prénom	6	0
numéro	16	0

Tableau.4 Taux d'apparition de certains substantifs dans notre corpus

On a noté dans notre corpus la présence d'éléments inattendus comme : "la noter" et "la notation"; Cela s'explique par une confusion [verbe /substantif] pour le terme "la note" ; mais le sens est même (erreur d'écriture).

Comme le montre la Capture .1, nous avons illustré deux exemples de ce genre d'erreurs (les dialogues s

<p>U4:-Vous pouvez me donner</p> <p>M4:-hum, pouvez vous me donner?</p> <p>U5:-je veux des notations des examens par exemple</p> <p>M5:-quel est votre institut, votre option et votre année ?</p> <p>U6:-électronique .communication</p> <p>M6:-voila les instituts qui existent : électronique, médecine, chimie, science de la terre</p> <p>U7:-électronique</p> <p>.....</p> <p>M3:-pourquoi vous questionnez?</p> <p>U4:-cé</p> <p>M4:-Ok. Vous avez dit "bon vous êtes ici pour ce qui concerne les résultats de l'examen final", encore plus.</p> <p>U5:-c'est mon droit de chercher ma noter</p> <p>M5:-quel est votre institut, votre option et votre année ?</p> <p>U6:-médecine</p> <p>M6:-voila les instituts qui existent : médecine, électronique, physique, science de la terre</p> <p>U7:-médecine</p> <p>.....</p>	<p>Ca ptu re.1 Err eur s d'é crit ure du mot "no te".</p>
---	---

3.4 Concept d'ouvertures et de clôture

Dans notre corpus (selon le tableau VI.4) les concepts d'ouverture sont très utilisés dans les dialogues, on observe par exemple presque dans tous les fichiers le mot "bonjour". Ce phénomène est sans doute dû au fait que le dialogue est sur l'initiative de la machine (appréciation de la machine à l'utilisateur). La machine interroge ce dernier concernant sa situation ("comment ça va ?"), Préférant entrer directement en dialogue collectif (comme entre deux être humain). Une fois le système s'est présenté, l'utilisateur s'imagine sans doute que le système est entrain d'entrer dans un dialogue coopératif.

Concept d'ouverture du dialogue	Concept de clôture du dialogue
Bonjour	Au revoir
Salut	A bientôt
Bonsoir	Bey
	Bye
	A plus tard
	A+
	Stop
	Finir
	Arrêter
	Partir
	Terminer
	Tchao
	abandonner

Tableau 4 Concept d'ouverture et de clôture de dialogues

La capture 2 présente des exemples d'ouverture de dialogues.

<i>M1:- salut, bonjour</i>
<i>U1:- je m'appelle amine.</i>
<i>M2:- eh, comme ça toujours vous ne dites pas bonjour !</i>
<i>U2:- donc, bonjour</i>
<i>M3:- comment ça va aujourd'hui?</i>
<i>U3:- très bien</i>
<i>M4:- c'est pas mal, très bien !</i>
<i>U4:- oui</i>
.....
<i>M1:- salut, bonjour</i>

U1:- *bonjour*
M2:- *comment ça va aujourd'hui?*
U2:- *bien*
M3:- *ah, c'est bon comme ça!*
U3:- *c'est bon*
.....
M:- *salut, bonjour*
U:- *allô.*
M:- *eh, comme ça toujours vous ne dites pas bonjour !*
U:- *excusez moi bonjour bonjour!*
M:- *comment ça va aujourd'hui?*
U:- *sa marche un peu*
.....

Capture 1. Exemples d'ouverture des dialogues.

Les concepts de clôture de dialogue sont très utilisés dans notre corpus, par exemple les mots (bye et au revoir) sont utilisés dans presque tous les dialogues, auquel "au revoir" est utilisé par environ 40% des usagers, "byby" (erreur introduite par les usagers) 20% ; "a bientôt" 10%, "bye" (10%)...etc. Ces pourcentages sont calculés par rapport au nombre total des fichiers recueillis, cela implique que plus de 90% des usagers ont accompli leur requête et quittant le dialogue satisfaits (un seul étudiant parmi dix n'a pas accompli sa requête). La capture IV.3 présente des exemples de clôture de dialogue.

U:- *byby (erreur de frappe)*

M:- *au revoir, merci pour votre collaboration, si vous voulez me parler fermez cette fenêtre puis ouvrez de nouveau.*

U:- *je veux ma note de tec 006*

M:- *j'ai dit fermez cette fenêtre puis ouvrez de nouveau, d'abord à partir de maintenant je ne te réponds pas.*

U:- *pourquoi ?*

M:-
.....

U:- *aller on se quitte, bye*

M:- *au revoir, merci pour votre collaboration, si vous voulez me parler fermez cette fenêtre puis ouvrez de nouveau.*

.....

U:- *merci beaucoup pour votre service et au revoir*

M:- *au revoir, merci pour votre collaboration, si vous voulez me parler fermez cette fenêtre puis ouvrez de nouveau.*

.....

Capture 2. Exemples de clôtures de dialogue.

3.5 La machine : une être familier?

En général l'utilisateur est bien adapté à notre système, dont parfois il met en doute les capacités et croit celui-ci remplacé par un être humain. Par exemple dans la capture IV.4, l'usager utilise le mot "idiot", car il prend la machine pour un être humain bête. Parfois l'usager assimile la machine à un nouvel ami dont il utilise des mots comme "ravi de votre rencontre". Cela montre bien que l'usager dialogue avec la machine de la même façon qu'il le ferait avec un être humain, essayant parfois de jauger les limites du système.

M :- *qu'est ce que vous pensez ?*

U :- *une simple idiote machine.*

.....

U :- *d'accord ravi de votre rencontre*

M :- *merci.*

.....

Capture 3 Exemples de ressemblance de la machine par un être humain.

Le tableau 6 suivant donne la liste de quelques pronoms du corpus ainsi que leur taux d'occurrences rencontré.

Pronoms	Homme (%)	Machine (%)
---------	-----------	-------------

Je	57.1	14.2
Tu	17.8	13.0
Il	0	7.3
Nous	0	0
Vous	21.4	63.6
On	3.5	1.7

Tableau 6 Pronoms personnels utilisés dans le corpus

Selon [Rouillard, 2000], jacquet dit que l'utilisation de "tu" signifie que les sujets s'adressent clairement au compère humain en situation de magicien d'oz. C'est à dire que le vouvoiement est remplacé par le tutoiement en signe " de reconnaissance " de coopérations entre hommes.

Dans notre expérience la machine vouvoie l'utilisateur dans la plupart des cas. Dans le tableau 6 le taux atteint "63.6%" de pronoms "vous" parmi le nombre total des pronoms énoncés par la machine, malgré que l'utilisateur tutoie le système dont le taux du pronom "tu" énoncé par l'utilisateur est presque égal au taux pour le pronom "vous". Ça ne signifie pas que le système ne tutoie pas l'utilisateur, dont on retrouve autant de "tu" : "13 %" des pronoms. Donc certains sujets pensent avoir affaire à un interlocuteur humain, ils dialoguent avec la machine comme avec un vrai partenaire.

4. Connaissances pour le modèle de dialogue

Le modèle de dialogue constitue le cœur du système de dialogue dont il définit la structure du dialogue entre l'utilisateur et la machine. Selon Caelen [Caelen, 1997], il définit et précise les deux composantes

- ↳ Les actes de dialogue autorisés.
- ↳ Les stratégies.

4.1 Les actes de dialogue

Le modèle des actes de dialogue précise ce que l'utilisateur a le droit ou non de dire et de faire. Il Précise le vocabulaire (lexique), la syntaxe et le type d'actes autorisés. Chaque interlocuteur peut énoncer des séquences de mots. Ces séquences correspondent à un ou plusieurs actes de langage. Un acte de dialogue doit donc être segmenté en actes de langage qui doit être reconnu avant d'être interprété. Un acte peut contenir plusieurs actions et réciproquement.

Les actes de dialogue décrivent les intentions et les besoins de celui qui parle (chercher l'information, présentation de l'individu, ...). Les concepts précisent à propos de quoi l'acte de dialogue est exprimé (l'institut, l'option, l'année, ...). Les arguments permettent d'instancier les valeurs des variables du discours (électronique, chimie, contrôle, ...).

Par exemple pour un usager dont un tour de parole signifie "je voudrais voir ma note sachant que je suis en troisième année contrôle de l'électronique". L'acte de dialogue est décomposé comme suit :

H : Chercher information + nom de l'institut + nom de l'option + nom de l'année d'étude.
(
Chercher information = (ma note).

Nom de l'institut = (électronique).

Nom de l'option = (contrôle).

Nom de l'année d'étude = (trois).

)

H : indique que c'est l'utilisateur qui parle, chercher information + nom de l'institut + nom de l'option + nom de l'année d'étude est l'acte de dialogue; électronique, contrôle, trois sont des valeurs.

Nos corpus	Expressif	Assertif	Directif	Promissif	Déclaratif	Tours de parole
1	0	4	1	1	12	18
2	0	3	7	1	17	28
3	0	2	2	1	7	12
4	0	1	2	1	14	18
5	0	2	1	1	8	12
6	3	1	2	1	14	22
7	2	1	2	1	13	19
8	1	2	1	1	7	13
9	1	1	1	1	7	12
10	2	3	3	1	13	22
Totale du corpus	9	20	22	10	112	

Tableau 7 Actes de langage utilisés dans le corpus

N° corpus	Expressif (%)	Assertif (%)	Directif (%)	Promissif (%)	Déclaratif (%)	Tours de parole
1	0	22,2	5,5	5,5	66,6	18
2	0	10,7	25	3,5	60,7	28
3	0	16,6	16,6	8,3	58,3	12
4	0	5,5	11,1	5,5	77,7	18
5	0	16,6	8,3	8,3	66,6	12
6	14,2	4,7	9,5	4,7	66,6	22
7	10,5	5,2	10,5	5,2	68,4	19
8	8,3	16,6	8,3	8,3	58,3	13
9	9,1	9,1	9,1	9,1	63,6	12
10	9,1	13,6	13,6	4,5	59,1	22
Moyenne (%)	5,2	11,5	12,7	5,7	64,7	

Tableau 8 Actes de langage en pourcentage

4.2 Stratégies de dialogue

Dans le dialogue entre les humains il existe des stratégies très sophistiquées pour atteindre un but. Par contre dans le dialogue entre un système et un locuteur humain la stratégie est limitée par l'état mental de l'utilisateur. Dans [Cherabit & al, 2004] nous définissons la stratégie de dialogue par la façon de faire la requête par un utilisateur pour atteindre son but. Par exemple, selon le corpus il existe des locuteurs qui veulent dialoguer, et à un certain niveau de dialogue le système demande à l'utilisateur de poser sa requête. Alors que d'autres posent directement la

requête. A ce moment, on conclut que la stratégie de dialogue diffère d'un locuteur à un autre. La stratégie de dialogue dans notre modèle est de deux types "stratégies inférentielles et non inférentielles" [Caelen, 1997].

4.2.1 Stratégies non inférentielles

Ce type de stratégie ne met pas en jeu la compréhension du but de l'interlocuteur. Le système agit sans tenir compte du but de l'utilisateur. Il existe deux types dans cette stratégie :

4.2.2 Stratégie directive

L'initiative reste toujours du côté de la machine, l'utilisateur doit répondre strictement aux questions ou exécuter les ordres demandés. Ce mode est utile pendant la phase d'ouverture et la phase de clôtures de dialogue (au début la machine prend l'initiative avec une réponse "*salut, bonjour*"), demande de confirmer une information, ou en cas d'incompréhension d'un énoncé.

4.2.3 Stratégie réactive

Dans ce mode le locuteur réagit le plus complètement possible aux échanges. Ce genre de stratégie se trouve lorsque la machine se satisfait de répondre à l'utilisateur et qui n'apporte aucun élément d'aide ou renseignement, comme avec la réponse suivante :

- *ah, c'est bon comme ça!*

Dans ce mode, la machine répond par un acte à l'initiative de l'utilisateur par échanges réactifs jusqu'à un accord commun.

4.2.4 Stratégies inférentielles

Ce type de stratégie passe par la recherche du but de l'interlocuteur et met en relation les connaissances supposées et partagées. Dans ce cas on atteint un but par :

- ✓ Coopération entre le système et le locuteur : le système doit fournir un maximum d'informations pour aider et orienter son interlocuteur par exemple : donner quelques informations nécessaires à l'utilisateur "*introduisez un nom des instituts suivants : électronique, chimie, mathématique.* ";
- ✓ Dirigée par les intentions : C'est une sorte de stratégie coopérative fondée intensivement sur le but. Cela consiste à comprendre et interpréter les intentions de son interlocuteur. Il s'agit pour la machine d'un processus inférentiel pour déduire les objectifs de l'utilisateur à travers la succession de ses actes relativement au contexte de la situation.
- ✓ Construction : le principe de ce mode est d'apporter des informations nouvelles dans le discours, dans le but de provoquer une rupture dialogique si possible enrichissante vis-à-vis des connaissances partagées.

Notre modèle utilise trois stratégies de ce mode : tout d'abord une stratégie permettant de faire débiter le dialogue ; c'est une stratégie directive. C'est en effet la machine qui prend l'initiative du dialogue et faire orienter l'utilisateur d'une manière à fournir des informations nécessaires à la poursuite de l'interaction et la résolution du problème. Puis des stratégies réactives et coopératives.

Stratégie coopérative : la machine doit apporter son aide à l'utilisateur. L'introduction de la requête va entraîner une question concernant l'état de l'utilisateur. Le dialogue se poursuit jusqu'à ce que l'objet de la requête soit clair et la machine propose à l'utilisateur d'introduire des informations.

Stratégie de dialogue	N	%
-----------------------	---	---

Directive	24	14
Coopérative	67	40
Constructive	41	24
Réactive	38	22

Tableau 9 Stratégies de dialogue dans le corpus

- 1 : Stratégie directive
- 2 : Stratégie réactive
- 3 : Stratégie constructive
- 4 : Stratégie coopérative

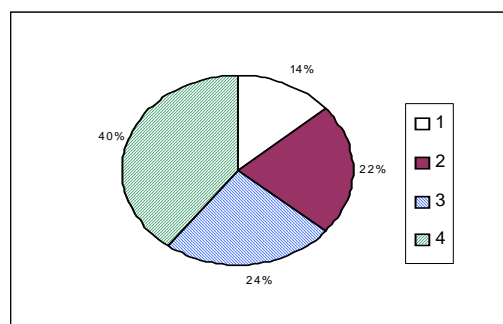


Figure 10 Stratégies de dialogue du système en pourcentage.

Selon les résultats obtenus (Figure VI.10), il semble que la stratégie la plus utilisée dans notre système a atteint (40%) pour la stratégie coopérative (c'est le cas d'un système de renseignement). Dans ce cas la machine essaye de coopérer avec l'utilisateur en de lui fournissant les informations nécessaires pour atteindre son but, suivie de la stratégie constructive (avec environ 24%) ici la machine tente de fournir des informations à l'utilisateur pour lui permettre de construire des plans en l'orientant vers la requête. La stratégie réactive atteint (environ 22%) et la stratégie directive (avec environ 14%).

5. Conclusion

Nous avons montré que sur les corpus acquis on pouvait lire les attitudes des usagers en situation de dialogue homme-machine réel. C'est ainsi que nous avons pu déterminer les actes de dialogue et les stratégies de dialogue utilisés dans notre système. Ces notions de stratégies de dialogue et les actes de dialogue utilisés pourront aider à mieux modéliser les dialogues pour mieux adapter le système et faciliter l'interaction entre l'homme et la machine à travers le dialogue. Donc pour mettre en œuvre des systèmes de dialogue homme-machine, il est nécessaire de travailler simultanément sur plusieurs aspects, notamment sur la compréhension de la langue naturelle et le comportement des usagers en situation de dialogue.

Après avoir analysé notre dialogue sémantiquement, nous allons passer dans ce qui suit à un autre niveau : **la pragmatique**. C'est dans cet esprit que nous allons nous intéresser à la représentation formelle d'un discours en vue de son interprétation, Nous passerons en revue les théories de représentations formelles. En nous basons sur la théorie de représentation de discours structurée (SDRT) nous donnerons quelques analyses et interprétation de dialogues.

VII. De la sémantique formelle à l'interface sémantico-pragmatique

7.1 Introduction

La sémantique formelle a été profondément affectée par le tournant discursif de la linguistique.

La sémantique de Montague en 1974, qui vise à déterminer les phrases sémantiquement bien formées, a évolué vers une sémantique du discours fondée sur le dynamisme du sens [Kamp, 1981, Heim, 1982].

7.2 Analyse discursive

L'analyse discursive cherche à déterminer des règles pour définir les discours *cohérents*

- **La cohérence** : c'est un ensemble qui exprime des idées qui s'accordent entre elles et non contradictoires
- **La Cohésion** : c'est un ensemble dont les parties sont unies et harmonisées

7.3 Les liens dans un discours

Les liens qui unissent les phrases dans un discours sont de trois ordres :

- les liens anaphoriques
- les liens temporels et spatiaux
- les liens rhétoriques

Les liens rhétoriques sont ceux qui structurent le discours par la fonction rhétorique d'une proposition par rapport aux autres.

7.4 Discourse representation theory (DRT) : (Kamp et Reyle (1993))

Selon cette théorie, la signification d'une phrase est représentée comme une forme logique qui sert à déterminer ses conditions de vérité. Elle ne prend en compte que les aspects sémantiques du discours, c'est-à-dire ceux qui sont liés à la signification des phrases.

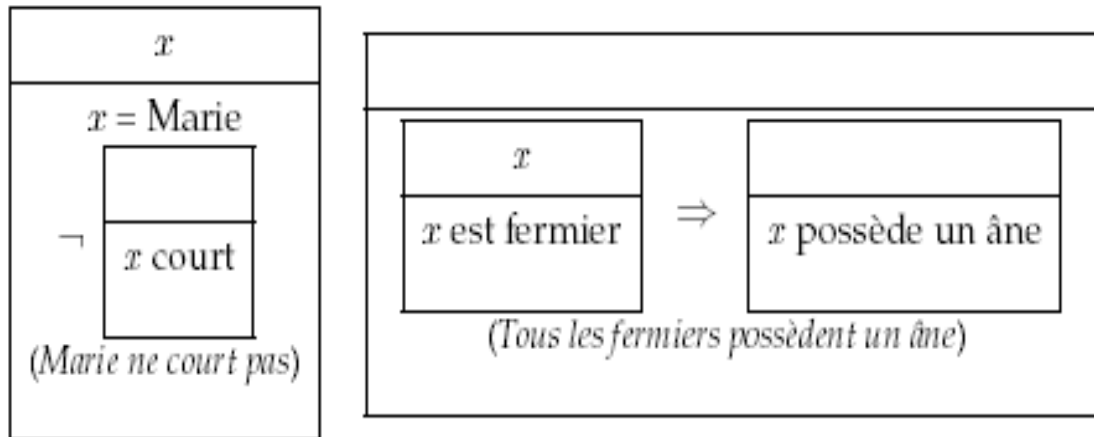
En DRT, un discours est une suite de phrases traitées successivement par des règles qui permettront de construire sa représentation sémantique. Ces représentations s'appellent DRS (*discourse representation structures*)

La DRT a pour mérite de fournir une procédure simple et structurée pour convertir une représentation syntaxique en représentation sémantique, elle a également tenté de résoudre certains problèmes comme l'anaphore et la présupposition. En revanche, en tant que théorie du discours, elle montre très rapidement de nombreuses limites. La plupart de ces processus agissent en effet à l'intérieur des phrases et il n'y a aucune prise en compte du contexte ni des relations qui existent entre les phrases. C'est pour tenter de résoudre certains de ces problèmes qu'Asher (1993), puis Asher & Lascarides (2003) ont proposé une version modifiée : la SDRT qui tente résoudre les problèmes d'interprétation incluant le contexte.

Représentation d'un exemple de DRS :

Exemples : « Marie ne court pas » « Tous les fermiers possèdent un âne »

La DRS pour ces deux phrases se présenterait ainsi



7.5 La SDRT : une théorie sémantico- pragmatique pour l'interprétation du discours

Cette théorie initiée par Asher (1993), puis Asher & Lascarides (2003) a introduit un ensemble de relations de cohérence, et des mécanismes logiques qui, sur la base du contenu propositionnel du discours et de certains principes de nature sémantico-pragmatique, révèlent l'articulation rhétorique du discours, c.-à-d. permettent d'inférer le découpage en segments et les relations qui les lient [Busquets et al 2001].

La SDRT est une théorie de représentation *segmentée* du discours en vue de son interprétation.

Elle rend compte de la structure hiérarchique du discours par le biais d'une structure de la représentation du discours *segmentée*.

Cette représentation s'appelle SDRS

Une SDRS est un constituant discursif complexe composé de constituants discursifs plus simples les DRS reliées entre elles par des relations de discours.

Les SDRS peuvent à leur tour être reliées entre elles par des relations de discours pour former des structures hiérarchiques

Un **discours** est constitué de segments reliés entre eux par des relations, ces relations sont appelées *relations rhétoriques*.

L'hypothèse de cohérence du discours se traduit du point de vue rhétorique par le principe que chaque segment du discours doit être relié, ou « attaché », à un autre segment du contexte par (au moins) une relation de discours : chaque segment doit être en rapport avec quelque chose dit précédemment.

7.6 Les relations rhétoriques (ou relations de cohérence ou encore relations structurelles):

- **Narration (a,b)** : l'événement décrit en *b* est une conséquence de l'événement décrit en *a*.
(1) *Max se leva. Jean le suivit.*

- **Explication (a,b)** : l'événement décrit en *b* explique pourquoi l'événement *a* s'est produit. (2) *Max est tombé. Jean l'a poussé*
- **Élaboration (a,b)** : l'événement *b* est une partie de *a*. (3) *Le conseil a construit le pont. L'architecte a fait les plans*
- **Arrière-plan (a,b)** : l'état décrit en *b* est la toile de fond (circonstances) dans laquelle l'événement *a* s'est produit. (4) *Max ouvrit la porte. La chambre était complètement noire.*
- **Résultat (a,b)** : l'événement décrit en *a* a causé l'événement ou l'état décrit en *b*. (5) *Max éteignit la lumière. La chambre était complètement noire.*

Les relations subordonnantes et coordonnantes

Tout discours possède une structure hiérarchique fondée sur une distinction entre

- Les relations subordonnantes, comme Élaboration, Explication ou Topique représentées par une flèche verticale dans une SDRS et notées Subord.
- Et les relations coordonnantes, comme Narration, Arrière-plan, Parallèle ou Contraste, représentées par une flèche horizontale dans une SDRS et notées Coord.

Les axiomes

A partir de la relation de narration; Lascarides et Asher (1993) introduisent deux axiomes :

- l'un (A₁) établissant la nécessité d'un topique commun, un thème commun,
- l'autre (A₂) précisant les effets temporels de cette relation de discours à savoir une relation de précedence temporelle entre les événements de deux constituants.

$\text{Narration}(\alpha, \beta) \rightarrow \exists \gamma (\text{Contingent}(\gamma) \& \gamma \Downarrow \alpha \& \gamma \Downarrow \beta) \& \neg(\alpha \Downarrow \beta) \& \neg(\beta \Downarrow \alpha)$

Contingent(φ) signifie que φ et $\neg\varphi$ sont tous les deux possibles.

$\alpha \Downarrow \beta$ signifie α que domine d'un point de vue discursif et qu'il constitue le topique de β .

Exemples

A titre d'exemples nous prenons les discours (1) et (2) dans lesquels chaque constituant a été étiqueté.

(1) Nabil a fait du jardinage (π^*). Il a tondu le gazon (π_1). Puis, il a taillé les arbres (π_2).

(2) Nabil a tondu le gazon (π_1). Puis, il a taillé les arbres (π_2).

La relation de narration (π_1, π_2) indique d'une part que les deux constituants π_1 et π_2 se réfèrent à un thème ou topique commun et d'autre part que l'ordre textuel de ces constituants reflète l'ordre temporel entre les événements décrits. L'exemple(1) ci après illustre cette relation de discours

L'axiome (A₁) nous permet de construire pour le discours (1) Un constituant complexe Π_0 regroupant π_1 et π_2 . La structure est présentée par la figure1 où les constituants π_1 et π_2 (respectivement associés aux phrases du discours (1)) sont reliés par la relation coordonnante Narration, grâce à l'adverbe *puis* dans π_2 .

π^* représente un topique, pour les deux constituants π_1 et π_2 . La relation subordonnante Élaboration peut être inférée entre π^* et Π_0 .

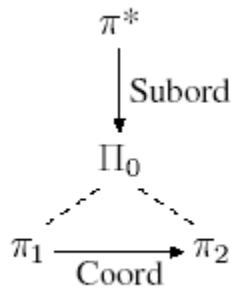


Figure 7.1 SDRT pour les exemples (1) et (2).

Règles

- **Le principe d'accessibilité des référents :** Si deux SDRS π_1 et π_2 sont reliées par une relation rhétorique $R(\pi_1, \pi_2)$, π_2 et ses sous-DRS accèdent aux référents (DRS-accessibles) de π_1 ; autrement dit On ne peut accéder qu'aux référents qui dominent le constituant courant, ou du constituant situé immédiatement à gauche.
Seuls les référents qui sont dans le site de rattachement ou dans les constituants qui le dominent peuvent faire l'objet d'une reprise anaphorique dans le nouveau constituant. Dans l'exemple de la SDRS de la figure 7.1. Si le nouveau constituant à rattacher est π_2 , et que son site de rattachement est π_1 alors les référents de π_1 et de π^* peuvent être les antécédents des expressions anaphoriques de π_2 .
- **Le principe de la poursuite du schéma discursif ou CDP.** Deux constituants reliés par une relation coordonnante, comme Narration, sont homogènes par rapport au constituant qui les domine. Par exemple, dans le discours (1), nous avons Élaboration (π^* , π_1) et Narration(π_1 , π_2). Avec le principe du CDP, nous obtenons Élaboration (π^* , π_2), ce qui correspond à l'interprétation qui a été représentée dans la figure 7.1.

Le schéma suivant (figure 7.2) illustre les définitions précédentes. La mise à jour d'une telle SDRS ne pourra attacher des relations de discours qu'aux nœuds indiqués (open and free). Le nœud π_{22} est le nœud en cours, et les nœuds mis en relief (π_0 π_{12}) sont ceux auxquels ce nœud est relié par une chaîne de relations de discours ou de topique.

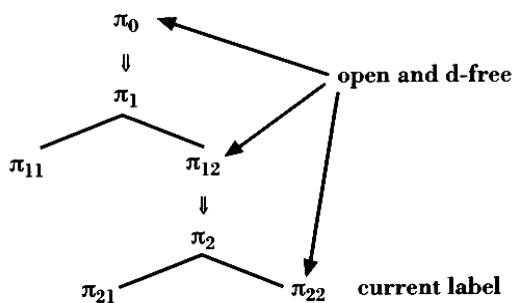


Figure 7.2: exemple de SDRS montrant le critère d'accessibilité. [Asher 1993]

Construction du discours). [Xuereb & Caelen, 2004].

Le processus de mise à jour de la structure du discours est un procédé incrémental :

- construire la DRS de la phrase courante ;
- intégrer ce segment dans le **contexte** de la SDRS déjà construite :
- décider quel référent discursif (segment de discours étiqueté π) peut constituer un site d'attachement.
- inférer la relation de discours qui relie cette nouvelle DRS à un site d'attachement disponible.

- mettre à jour la SDRS résultante : résoudre certaines sous-spécifications, introduire de nouveaux segments complexes.

En cas d'ambiguïté de rattachement, le principe de la Cohérence Maximale du Discours (MDC) permet de choisir l'interprétation préférée.

En cas d'ambiguïté dans le choix des relations rhétoriques et des sites d'attachement, on privilégie l'augmentation de la cohérence. Il faut donc, dans la construction de la structure, maximiser cette cohérence.

- on maximise le nombre et la qualité des relations rhétoriques.
- on minimise les sous-spécifications (on augmente le nombre de résolutions).

Afin de rendre compte des interactions entre les segments et la structure globale et pragmatique du discours nous présentons un exemple de la construction de la SDRS pour l'énoncé suivant discours (4).

(4) Nabil quitta Alger. Il voyagea à Béjaia. Là il alla à la plage avec Mahmoud.

La première case contient les *référents* du discours (4) : e_1 e_2 e_3 sont des référents, temporels, donc des référents d'événements, n représente le moment de parole.

La seconde case contient les *conditions* de la DRS : elles spécifient les référents du discours et définissent les relations entre eux.

Plus formellement, une DRS K est définie par la donnée par la donnée d'un couple

$x, y, z, t, u, v, w, e_1, e_2, e_3$
Nabil(x)
Alger(y)
e_1 : quitter(x, y)
$e_1 < n$
$z = x$
Béjaia(t)
e_2 : voyager-à(z, t)
$e_2 < n$
$e_1 < e_2$
$u = x$
Mahmoud(v)
Plage(w)
Aller-à($u \oplus v, w$)
$e_3 < n$
$e_2 < e_3$

Figure 7.3 construction

de SDRS exemple (4)

7.7 Les relations dialogiques.

rhétoriques

Un dialogue est constitué d'interventions procédant de différents tours de parole. Il peut y avoir dans ces interventions, des séquences monologiques ou dialogiques. Nous examinons maintenant les relations rhétoriques (RR).

Dans le cadre du dialogue finalisé les relations rhétoriques dialogiques peuvent être considérées comme formant les familles de relations suivantes.

- **1 L'axe épistémique et son subordonné (sémantiquement)**
 - couples question-réponse notés **QAP** (Question Answer Pair).
 - **Les questions subordonnées notées Q-Sub**
 - **Les élaborations de connaissances notées Élab**
- **2 axe actionnel et son subordonné**
 - **Les délégations d'action notées RAP**
 - **Les élaborations de plan notées P-Élab**
 - **Les élaborations de question Q-Élab**
 - **Les élaborations de but R-Élab**
 - **3 les axes divergents : incidences et répliques** Les incidences notées **I**
 - **.4 Les continuations notées C**
 - **Les répliques notées R**
- **4 Les continuations notées C**
- **5 topique**

7.8 Analyse du dialogue

A titre d'exemple nous faisons une analyse pour le couple $F_A^{FS} \rightarrow F_B$ et le **Couple $F_A^F \rightarrow F_B$** . F_A^{FS} est une demande de renseignement (question) sur l'arrière-plan ou la situation. Le contenu propositionnel de la question est de nature épistémique et non de nature actionnelle. Mais la réponse peut avoir un contenu actionnel (répondre par l'action). La question peut être déniée par B. F^{FS} crée un double engagement de A et de B. On suppose que A pose la question et B fournit une certaine réponse dans la visée de A.

a- **Couple $F_A^{FS} \rightarrow F_B$**

F_A^{FS} est une demande de renseignement (question) sur l'arrière-plan ou la situation. Le contenu propositionnel de la question est de nature épistémique et non de nature actionnelle. Mais la réponse peut avoir un contenu actionnel (répondre par l'action). La question peut être déniée par B. F^{FS} crée un double engagement de A et de B. On suppose que A pose la question et B fournit une certaine réponse dans la visée de A. Exemples :

F^{FS}_A : Puis je avoir ma note de TEC 588 ?

Réponses possibles

		Rd
F^S_B :	- Vous avez 12/20.	[QAP]
	- Vous avez au dessous de la moyenne	[PQAP]
	- Vous êtes noté par un C.	[IQAP]
	- Les notes sont sur le site de la fac.	[P-Élab]
	- Elle compte pour votre passage ?	[Q-SUB]
	- Je n'ai pas encore corrigé	[QAP]
	- J'ai égaré votre copie!	[R]
F^{FS}_B :	- Vous êtes bien inscrit chez moi ?	[Q-SUB : vérification]
	- Vous êtes sur de vouloir la connaître?	[Élabq]
	- C'est pour votre moyenne ?	[Q-SUB: clarification]
	- Avez-vous bien travaillé ?	[Q-SUB : incidence]
F^F_B :	- Allez voir l'affichage	[QAP]
	- Consulter le site WEB	[IQAP]
	- Vérifiez d'abord si vous êtes sur ma liste	[R-Élab]
	- Repassez à l'oral	[PQAP]
	- Demandez à votre assistant	[R-Élab]
F^A_B :	- Venez! je vais vous la montrer	[QAP]
	- Attendez que je regarde ma fiche	[PQAP]
	- Je n'ai pas de temps à vous consacrer	[R]
F^D_B :	- Vous devez d'abord me donner votre nom	[R-Élab]
F^P_B :	- Vous avez la même note que votre binôme	[P-Élab]

b- Couple F^F_A → F_B :

F^F_A est une demande d'action. A propose une action à B et en partage les effets. Cet acte crée un but qui engage A et B. L'acte s'interprète différemment selon la stratégie de A (par exemple en stratégie directive, A veut atteindre son but, en stratégie coopérative ce sera plutôt une invite), mais quoiqu'il en soit cet acte pose un but en arrière-plan. Exemples :

F^F_A : Pouvez vous me donner ma note de TEC 588 ?

Réponses possibles

		Rd
F^S_B :	- Je peux vous donner la fourchette.	[PRAP]
	- Oui si vous êtes sage.	[IRAP]
	- Je ne suis pas disponible.	[R]
F^{FS}_B :	- C'est pour calculer votre moyenne ?	[Élabq]
	- C'est bien votre note que vous demandez?	[Q-SUB: clarification]
	- Etes vous un doublant ?	[Q-SUB : vérification]
	- Vous croyez que j'ai du temps à perdre ?	[R]
F^F_B :	- Vous pouvez consulter le site de la fac	[R-Élab]
	- Faites plus d'efforts	[R]
F^A_B :	- Voilà votre note	[RAP]
	- Venez! je vous montre la liste	[PRAP]
F^D_B :	- Vous devez vous identifier	[R-Élab]
F^P_B :	- Vous pouvez avoir votre note à partir du site	[P-Élab]

7.9 Conclusion

Ce chapitre aura montré que l'application de La SDRT au discours est plus ou moins maîtrisée mais la représentation du dialogue par le SDRT est un sujet très ardu et met en jeu les objets du message linguistique à savoir les référents, les relations de discours et les actes de parole et nous l'avons peut être survolé de manière un peu générale. Néanmoins, en nous basant sur les travaux de Caelen et Xuereb [Caelen et al 2004] nous sommes parvenus à l'appliquer à notre corpus et capturer les différents niveaux de structure.

Le travail présenté ici n'est qu'une étape dans la description des composantes du dialogue homme machine. Elle doit maintenant être étendue à d'autres corpus et d'autres types d'interactions verbales. L'intérêt grandissant pour les études de corpus dans le domaine sémantique promet l'exploitation et la poursuite du travail réalisé ici, la SDRT offre un cadre de modélisation riche pour le dialogue homme machine et reste dans nos perspectives prochaines de développement.

Conclusion générale et perspectives

Dans ce travail, nous avons mis en place trois systèmes de communication personne machine. Dans le premier système, nous avons étudié les phénomènes mis en jeu lors de la gestion d'événements multimodaux autour d'une application scolaire. Il s'agit de remplir un document en utilisant soit le clavier ou le geste souris. Dans le deuxième travail, nous avons introduit dans la chaîne, l'étape de reconnaissance de la parole et celui de la synthèse. Nous avons alors montré que les différents étages de ce système nécessitent une certaine adaptation. Nous avons remarqué qu'en changeant d'application (nous avons mis en place une application graphique), les interactions entre les étages sont différentes. Un modèle adaptatif est proposé. Il est bien clair que les systèmes multimodaux sont des systèmes interactifs qui autorisent l'interaction à travers des modes et des médias divers. Actuellement, il existe des systèmes plus au moins évolués, mais ils restent encore limités car très peu sont les expériences qui permettraient de dégager des règles et des concepts stables pour la communication homme-machine de manière globale et multimodale en particulier. Nous avons donc adopté une démarche expérimentale pour mieux cerner les problèmes spécifiques à ce type de communication. Nous avons implémenté le module qui traite les gestes de désignation avec la souris et le langage naturel écrit. Il s'agit maintenant d'intégrer techniquement les diverses modalités qui participent à l'interaction. Pour cela, nous avons combiné la représentation en CMR du MMI2 et celle de la vue en couche du modèle de Seeheim. Le langage naturel ainsi que les gestes sont traités seuls avant de subir une fusion pour compléter le tableau d'analyse. Nous avons montré que le traitement du langage naturel fait appel à une analyse lexicale, syntaxique et sémantique pour tirer l'information pertinente. Nous avons ensuite traité les cas de la multi modalité synergique ou composée. Dans la première expérience les deux modalités ne sont pas faciles à exploiter d'où le montage de la deuxième expérience dans laquelle on a introduit la modalité parole. La gestion des événements multimodaux montrent clairement l'importance des contraintes temporelles. L'usage des références ou des coréférences parole et geste permet des solutions aux problèmes d'ellipses, d'anaphores ou d'interprétation des erreurs liées à l'interaction avec la RAP, le TTS et le MD (module de dialogue).

Ces deux expériences nous ont servi pour élaborer un nouveau système dont on peut faire évoluer les connaissances et travaille en bilingue, la langue arabe et la langue française. Nous avons proposé dans ce travail un modèle de dialogue homme-machine finalisé. Ce modèle prend place dans les recherches en communication visant à étudier les interactions entre l'homme et la machine. Les deux agents (l'agent humain et l'agent machine) ont à gérer conjointement une tâche en associant à des actes de communication, des actes informatiques (l'accès à une base de données et la recherche d'informations) pour satisfaire l'utilisateur en lui faisant atteindre son but. Notre modèle offre la souplesse nécessaire à la gestion de l'interaction, on voit ses capacités quand il convient d'exprimer naturellement les problèmes des usagers, de définir conjointement la tâche à effectuer, ou encore d'utiliser les erreurs et incompréhensions pour éviter les interruptions. Dans notre système, nous avons donc mis en place un modèle qui peut soit satisfaire le locuteur pour atteindre son but qui est "chercher une information", soit faire évoluer l'état de la machine en lui faisant un apprentissage de nouvelles informations.

L'enjeu principal était de construire un terrain commun où l'utilisateur et la machine puissent communiquer. Le langage naturel est la modalité idéale pour la construction de ce terrain commun. Ce système a permis l'enregistrement de corpus dialogiques via le réseau intranet.

L'analyse de ce corpus, et plus précisément l'analyse de l'enchaînement conversationnel et des négociations sur le sens en contexte des énoncés, nous a permis d'observer les attitudes de l'utilisateur mis dans une situation d'interaction avec la machine.

Une perspective de notre travail est l'amélioration de l'apprentissage de nouvelles situations apparaissant au cours de l'évolution de la communication et qui fournirait une acquisition automatique des connaissances nouvelles à la machine. Dans ce cas, on donne au système la possibilité d'apprendre et d'augmenter sa base de connaissances à chaque conversation. Il fallait pour cela pouvoir représenter le dialogue sous une forme structurée et définir un espace de représentation. La SDRT est une méthode qui pourrait convenir, c'est ce que nous avons essayé d'appliquer au corpus obtenu lors de notre dernière expérience. Le but de la modélisation par la SDRT est de construire la structure logique du dialogue. C'est cette structure logique qui guidera l'interprétation pragmatique en contraignant le processus de résolution.

Les relations dialogiques sont caractérisées selon :

- leur critère de reconnaissance
- leur influence spécifique sur la structure donc sur la sémantique de l'interprétation.

Dans le contexte du traitement de la parole spontanée, les résultats de l'analyse syntaxique et sémantique (après reconnaissance automatique de la parole) ne sont pas suffisamment fins pour exploiter des indices linguistiques élaborés.

Un dialogue est constitué d'interventions procédant de différents tours de parole.

Nous avons montré dans le chapitre VII que si l'application de La SDRT au discours est plus ou moins maîtrisée il n'en est pas de même dans sa représentation du dialogue ; en effet c'est un sujet très ardu car il met en jeu les objets du message linguistique à savoir les référents, les relations de discours et les actes de parole. Néanmoins, en nous basant sur les travaux de Caelen et Xuereb [Caelen et al 2004] nous sommes parvenus à l'appliquer à notre corpus et capturer les différents niveaux de structure.

Le travail présenté ici n'est qu'une étape dans la description des composantes du dialogue homme machine. Elle doit maintenant être étendue à d'autres corpus et d'autres types d'interactions verbales et même multimodales.

L'intérêt grandissant pour les études de corpus dans le domaine sémantique promet l'exploitation et la poursuite du travail réalisé ici, la SDRT offre un cadre de modélisation riche pour le dialogue homme machine et reste dans nos perspectives prochaines de développement.

Références bibliographiques

- [Allen & all, 1980] J F. Allen et C. R Perrault. *A Plan-Based Analysis of Indirect Speech Acts*. American Journal of Computational Linguistics, Volume 6, Number 3-4, July-December 1980.
- [Allen & all, 1980] J. Allen, CR Perrault, *Analysing Intention in Utterances*. In *Artificial Intelligence*, Vol 15 Pp 148-178 (1980).
- [Amsili & all] P. Amsili, J.Busquet, E.Vallduvi. *Etude comparative des connecteurs et des marqueurs discursifs dans le cadre d'une sémantique dynamique du discours*. Projet PICS France-Catalogue
- [Asher & all, 1995] N. Asher, M. Aurnague, M. Bras, P.Sablavyrolles, L. Vieu, *De l'espace-temps dans l'analyse du discours*. Sémiothiques, n°9, décembre 1995.
- [Asher, 1996] N. Asher, *L'interface pragmatique-sémantique et l'interprétation du discours*, Langages, Année 1996, Volume 30, Numéro 123 ; p. 30 – 50.
- [Baker, 2000] M. Baker. *Explication, Argumentation et Négociation : analyse d'un corpus de dialogues en langue naturelle écrite dans le domaine de la médecine*. Psychologie de l'Interaction, N° 9-10, 179-210.
- [Bateman, 1995] J A. Bateman, R. Henschel, and F. Rinaldi. *The generalized upper model 2.0. Technical report*, GMD/Institut für Integrierte Publikations- und Informationssysteme, Darmstadt, Germany, (1995).
- [Baudel] T. Baudel, *Spécificités de l'interaction gestuelle dans un environnement multimodal*. Laboratoire de Recherche en Informatique URA 410 du CNRS. Université de Paris-Sud.
- [Béguin et AL, 2004] Pascal Béguin et Yves Clot. *L'action située dans le développement de l'activité*. CNAM, Paris. @ctivités, I (2), 27-49. <http://www.activites.org/v1n2/beguिन.fr.pdf>
- [Bellik & All]Y. Bellik & D. Teil, SPECIMEN : *un outil pour la spécification des interfaces multimodales*. limsi-cnrs.
- [Bellik]Y. Bellik, *La composante temporelle dans les interfaces multimodales*. LIMSI-CNRS.
- [Beust, 1998] P. Beust. *Contribution a un modèle interactionniste du sens, amorce d'une compétence interprétative pour les machines*. Thèse de doctorat. Université De Caen.1998.
- [Bilange,1991] Bilange, *"Modélisation du Dialogue Oral Finalisé Personne-Machine par une ApprocheStructurale."*, Thèse d'Université en Informatique de Rennes (décembre 1991).
- [Bousquet-vernhettes, 2002] c. Bousquet-vernhettes. *Compréhension robuste de la parole spontanée dans le dialogue oral homme-machine – Décodage conceptuel stochastique*. Thèse de doctorat de l'université TOULOUSE III.

- [Bracops, 2006] M. Bracops, *"Introduction à la pragmatique, Les théories fondatrices : actes de langage, pragmatique cognitive, pragmatique intégrée."* www.deboeck.com
- [Busquets et al, 2001] J. Busquets, L.Vieu et N. Asher. *La SDRT : une approche de la cohérence du discours dans la tradition de la sémantique dynamique.* Verbum, vol 23 n)1, p73-101, 2001.
- [Caelen, 1997] J. Caelen. *Dialogue homme-machine et interaction verbale : une logique dialogique.* Marseille-Luminy,1997.
- [Caelen, 2002] J. Caelen. Modèles formels du dialogue. Interaction, Intelligence. Actes des 2e Assises nationales du GdR I3. Cepaduès Editions, (décembre 2002).
- [Caelen, 2003] J. Caelen. Stratégies de dialogue. MFI- mai 2003.
- [Caelen] J. Caelen, *Dialogue homme-machine et recherche d'information.*
- [Caelen & al] J. Caelen, AL Frechet, *Attitudes cognitives et actes de langage en situation de communication homme-machine.*
- [Caelen & al, 1998] J. Caelen, L Imberdis. *Génération d'actes illocutoires pour le dialogue* ACTES DE TALN 1998, PARIS, 10-11-12 JUIN 1998
- [Carthy, 1987] J. Mc. Carthy. Generality in artificial intelligence. Stanford University, 1987.
- [Cherabit & al, 2004] N. Cherabit, R. Djeradi et A. Djeradi. Stratégie de dialogue homme-machine en langue arabe. CNIE'04, USTO. 22-23 novembre 2004.
- [Cherabit & al, 2005 a] N. Cherabit, R.Djeradi. A. Djeradi; Modélisation Du Dialogue Homme Machine en Langue Arabe. SETIT 2005, MARCH 27-31, 2005 – TUNISIA.
- [Chereabit & al, 2005 b] CH. Noureddine, DJ. Rachida, DJ. Amar. Modélisation Du Dialogue Homme Machine bilingue. EMP 2005.
- [Clerc] D. CLERC, *la théorie des actes du langage.* damien.clerc@bluewin.ch
- [Curro & al, 1998] V. Curro, A d'attri et L Tarrantiro. Man-Machine interfaces, to medical information systems. IEEE engineering in medicine. 10th annual international conference. 1988
- [Djeradi & al, 2007] R.Djeradi, A.Djeradi & J Caelen. Modelling of the Human-Machine dialogue in dual language. ICA . Madrid 07.
- [Djeradi & al, 2001] R.Djeradi, A.Djeradi & J Caelen. Understanding of enouncing by the concept way. Acoustic's 08. Paris.
- [Dumazeau, 2001] C. Dumazeau. *Analyse empirique des communications distantes dans le cadre du contrôle aérien.* Mémoire de DEA d'ergonomie, CENA Berlin, Août 2001.
- [Dutoit] Thierry Dutoit, Article : *Synthèse de la parole*, MULTITEL – Département Synthèse de la parole (www.multitel.be/TTS). *How Text-to-Speech Works, Microsoft speech documentation. A Short Introduction to Text-to-Speech Synthesis*, TTS research team, TCTS Lab.
- [Fouquet, 2002] Y. Fouquet. *Un modèle de dialogue par les attentes du locuteur.* TALN Nancy, 24-27juin 2002.

- [Fougères, 2003] A.-J. Fougères. *Une architecture cognitive d'agents communicants dans des systèmes d'information complexes*. MFI- mai 2003.
- [Guinard] D. Guinard, P. deAlmeida, M. Ritz, *Le Service Counter System Toolkit, Une boîte à outils pour la construction d'interfaces multimodales*. Département d'Informatique, Université de Fribourg.
- [Jerome, 2003] A. Jerome, *Une architecture cognitive d'agents communicants dans des systèmes d'information complexes*. MFI- mai 2003.
- [Joab & al, 1999] M. Joab et C. Rossari. *Analyse automatique de séquences explicatives fondée sur le modèle genevois d'analyse du discours*. Psychologie de l'Interaction (9-10), 1999.
- [Jackiewicz & al, 2003] A. Jackiewicz, J-L. Minel. *L'identification des structures discursives engendrées par les cadres organisationnels*. TALN 2003.
- [Kerbrat-Orecchioni, 2003] C. Kerbrat-Orecchioni. *Les genres de l'oral : Types d'interactions et types d'activités*. ICAR, Université Lumière Lyon 2 Institut Universitaire de France.
- [Kerbrat-Orecchioni, 2001] C. Kerbrat-Orecchioni. *Les actes de langage dans le discours. Théorie et fonctionnement*. Nathan, 2001.
- [Latapy & al, 2003] M.Latapy, P. Dagorret et P. Lopisteguy. *La coordination intra – processus : ses interactions verbales*. Actes des Secondes Journées Francophones Cépadués, 2003.
- [Lehuen, 1997] J. Lehuen. *Un modèle de dialogue dynamique et générique intégrant l'acquisition de sa compétence linguistique Le système COALA*. Thèse de Doctorat Université De Caen. 1997.
- [Lehuen, 1999] J. Lehuen et T. Lemeunier. *Un modèle de génération des intentions de communications pour le dialogue homme-machine*. AT, TALN juillet 1999.
- [Lemeunier, 2000 a] T. Lemeunier. *Modélisation dialogique des intentions de communication en dialogue homme-machine*. Lyon septembre 2000.
- [Lemeunier, 2000 b] T. Lemeunier. *L'intentionnalité communicative dans le dialogue homme-machine en langue naturelle*. Thèse de doctorat de l'université du Maine décembre 2000.
- [Maudet et all, 2004] N. Maudet, P. Mullery et L. Prévoty. *Tableaux conversationnels en SDRT*. Workshop SDRT, TALN-04, Fès, 22 avril 2004.
- [Mermet, 2003] B. Mermet. *Collaboration dynamique dans un SMA*. MFI- mai 2003.
- [Moeschler,1989] J. Moeschler. *Modélisation du dialogue. Représentation de l'inférence argumentative*. Paris: Hermès. 1989.
- [Moeschler,2004] J. Moeschler. *Dialogue et causalité : force causale, actes de langage et enchaînement*; Département de linguistique Université de Genève,
<Jacques.Moeschler@lettres.unige.ch >
- [Microsoft] *Microsoft Speech API 4.0 HELPS: ActiveX controls, Microsoft*.
- [Nicolle, 2001] A. Nicolle. *Quels modèles de la langue et de l'activité langagière pour le dialogue personnes/machines ?*. Université de Caen cedex. Décembre 2001.
- [Nigay] L. Nigay, J. Goutaz, *Espace, problème, fusion et parallélisme dans les interfaces multimodales*. LIG-IMAG.

- [Nigay] L. Nigay, *Interaction Homme-Machine, Architecture logicielle des systèmes interactifs*.
- [Pasquier et al, 2002] P. Pasquier et B. Chaib-Draa. *Cohérence et conversations entre agents : vers un modèle basé sur la consonance cognitive*. 1^{re} soumission à JFIADSMA'02. juin 2002.
- [Patry, 1997] G. Patry. *Evaluation du Contexte de Travail par l'utilisateur*. IHM'99, Eds. J. Nanard & P. Girard, pp 118-125 novembre 1999.
- [Pesty & al 2003] S. Pesty and G. Chicoisne. *De l'interaction à la conversation entre agents. Une application à un agent de e-commerce d'une boutique virtuelle 3D*. IHM Leibniz-IMAG, mars 2003.
- [Prévot, 2004] Laurent PREVOT, *Structures sémantiques et pragmatiques pour la modélisation de la cohérence dans des dialogues finalisés*. THESE DE DPOCTORAT. Spécialité : Informatique – Intelligence Artificielle. UNIVERSIT'É PAUL SABATIER - TOULOUSE III
- [Prince, 1996] V. Prince *Vers Une Informatique Cognitive dans les Organisations : le rôle Central du Langage*. Editions Masson, Paris. 200 pages (1996).
- [Rosset & al, 2001] S. Rosset et L. Lamel. *Gestionnaire de dialogue pour un système d'informations à reconnaissance vocale*. TALN 2001, Tours, 2-5 juillet 2001.
- [Rouillard, 1998] J. Rouillard, *Contribution à l'étude du dialogue homme-machine à travers le web*. (RECITAL'98), Le Mans, 8-9 septembre 1998.
- [Rouillard & al, 1999] J. Rouillard and J. Caelen. HALPIN: *A multimodal and conversational system for information seeking on the World Wide Web*. Presented at ESCA ETRW workshop: Accessing information in spoken audio, Cambridge (UK), Avril 1999.
- [Rouillard, 2000] J. Rouillard. *hyperdialogue sur internet : le système halpin*. Thèse de doctorat, université Joseph Fourier – Grenoble 1. janvier 2000.
- [Rouillard, 2001] J. Rouillard. *Dialogue et Multimodalité*. CLIPS IMAG, Revue d'Interaction homme-machine Vol 2 N°1, 2001.
- [Sabah, 1997] G. Sabah. *Intelligence artificielle et sciences cognitives. LIMSI — Groupe Langage et Cognition, Informations In Cognito 1997*.
- [Sabah] G. Sabah, *Dialogue et sciences cognitives*.
- [Sadek, 1996] D. Sadek. *Interactions utilisateurs-services : de l'ergonomie des interfaces à l'agent intelligent dialoguant*. 1996.
- [Savall, 2003] M. Savall. *Une Architecture D'agents Pour La Simulation Le Modèle Yamam Et Sa Plate-Forme Phoenix*. Thèse Doctorat De L'insa De Rouen. Juin 2003
- [Schneider, 1996] D. K. Schneider. *Modélisation de la démarche du décideur politique dans la perspective de l'intelligence artificielle*. Thèse de doctorat de l'Université de Genève, Genève, 1996.
- [Secarea & al, 1970] V. V. Secarea, Jr. *Beyond. Knobs and dials: toward and intentional model of man-machine interaction*. IEEE page 736-769. 1990.
- [Suso Lopez 2003] Javier Suso Lopez. *La linguistique Française au XXe siècle*. Université de Grenade

- [Thevenin, 2001]** D. Thevenin. *Adaptation en Interaction homme-machine : le cas de la Plasticité*. Thèse de doctorat. Joëlle Coutaz. Décembre 2001.
- [Thull, 1993]** B.Thull. *Man-machine interaction in critical care settings. IEEE engineering in medicine and biology. December 1993*.
- [Torasso & al, 1989]** P. Torasso, L. Conole, P. Terenziani, et G. Molino. *man-machine interaction in deep diagnostic systems*. IEE ingeneering in medicine, 11th annual international conference.1989
- [Van Elslande, 1992]** P. Van Elslande. *Les erreurs d'interprétation en conduite automobile: mauvaise catégorisation ou activation erronée de schémas*. Intellectica, 1992/3, 15, pp. 125-149.
- [Tamen 1996]** Z. Tamen. *Mise au point d'un système interactif Homme-Machine multimodal autour d'une application de gestion de la scolarité*. Thèse magister 1996. USTHB.
- [Villaseñor-Pineda, 1999]**L. Villaseñor-Pineda. *Contribution à l'apprentissage dans le dialogue homme-machine*. Thèse de doctorat de l'université Joseph Fourier. Février 1999.
- [Xuereb & al 2004]** A. Xuereb, J. Caelen. *Un modèle d'interprétation pragmatique dialogue homme-machine basé sur la SDRT*. CLIPS-IMAG.
- [Xuereb 2004]** A. Xuereb. *Pragmatique du dialogue homme-machine*. Master 2004. CLIPS-IMAG.