CRIN *Université de Nancy I* Supelec-Metz

Laurent ROMARY

VERS LA DÉFINITION D'UN MODELE COGNITIF AUTOUR DE LA REPRÉSENTATION DU TEMPS DANS UN SYSTEME DE DIALOGUE HOMME-MACHINE.

Thèse soutenue publiquement le 14 avril 1989.

Président :	Jean Paul HATON.
Directeur de thèse :	Jean-Marie PIERREL.
Rapporteurs :	Mario BORILLO.
	Pierre LESCANNE.
Examinateurs :	Marion CRÉHANGE.
	Daniel KAYSER.
	Claude LHERMITTE.

CRIN *Université de Nancy I* Supelec-Metz

Laurent ROMARY

VERS LA DÉFINITION D'UN MODELE COGNITIF AUTOUR DE LA REPRÉSENTATION DU TEMPS DANS UN SYSTEME DE DIALOGUE HOMME-MACHINE.

Thèse soutenue publiquement le 14 avril 1989.

Président :	Jean Paul HATON.
Directeur de thèse :	Jean-Marie PIERREL.
Rapporteurs :	Mario BORILLO.
	Pierre LESCANNE.
Examinateurs :	Marion CRÉHANGE.
	Daniel KAYSER.
	Claude LHERMITTE.

Le présent travail n'aurait jamais pu voir le jour sans le soutien que m'ont procuré de nombreuses personnes au Crin et à Supelec-Metz. Tout d'abord, les idées que j'ai développé ont énormément profité des débats au sein de l'équipe dialogue du Crin ainsi que des échanges nombreux et productifs que j'ai pu avoir personnellement avec Jean Marie Pierrel, mon directeur de thèse. Je l'en remercie particulièrement.

De manière générale, l'ambiance ouverte qui régne au Crin donne un attrait particulier à ce laboratoire qui, malgré son importance toujours croissante, à su garder une homogénéité dans laquelle je compte m'insérer pleinement pour mes recherches à venir. C'est ainsi que je remercie Jean Paul Haton pour avoir suivi constamment mes travaux malgré les lourdes taches qui lui incombent et d'avoir bien voulu participer à mon jury.

Je tiens aussi à remercier ici MM. Lhermitte et Renouard qui, à Supelec Metz m'ont procuré un environnement de recherche propice à un travail rapide et productif, mais aussi des encouragements nombreux et réguliers qui n'ont pu que me motiver. La présence de Claude Lhermitte à mon jury est un signe de ce soutien constant.

J'exprime toute ma gratitude aux différents membres de mon jury, qui, suivant leurs disponibilités propres ont jugé mon travail en fonction du recul qu'ils possèdent vis à vis du domaine touché par ces recherches :

- Mario Borillo et Pierre Lescanne m'ont ainsi fait prendre conscience du travail encore a effectuer de par les différentes remarques qu'ils m'ont fait dans le cadre de leur rapports.
- Daniel Kayser m'a tout spécialement fait profité de son expérience par les nombreuses questions qu'il a pris le temps de me transmettre au sujet de cette thèse.
-Marion Créhange a montré l'intérêt qu'elle portait à mes recherches et c'est un plaisir de la voir figurer dans mon jury.

Enfin, j'adresse tout mes remerciements aux lecteurs et lectrices, parmi lesquelles Andrée et Laurence qui ont apporté de nombreuses modifications à la forme de cet exposé. Il en a gagné une lisibilité que seul, je n'aurais pu lui donner.

TABLE DES MATIERES.

0 INTRODUCTION.

Le travail présenté dans cette thèse concerne un ensemble de réflexions liées à des problèmes qui se posent dans le cadre d'un Dialogue Oral Homme-Machine. Les difficultés propres à ce domaine proviennent de la nécessité de maîtriser en parallèle de nombreuses informations de sortes différentes, pour effectuer des traitements linguistiques qui vont jusqu'à la gestion des dialogues, mais aussi pour représenter de multiples connaissances relatives au domaine de la tâche et raisonner efficacement sur ces connaissances.

L'approche que nous avons choisie est à la fois pragmatique et théorique. Pragmatique, car nous avons cherché dans un premier temps à résoudre un problème particulier dans le cadre d'une architecture de système existant dans notre équipe de travail au CRIN à Nancy. Théorique, car nous proposons en définitive un modèle de représentation des connaissances temporelles qui vise à reconsidérer les phénomènes cognitifs sous-jacents à l'activité de dialogue. Plus spécialement, nous désirons soutenir ici la thèse que les informations manipulées dans un dialogue homme-machine ne doivent pas être disséminées en autant de modules indépendants mais intégrées pour faciliter les échanges dans une architecture cohérente tant au niveau de la langue qu'au niveau des concepts manipulés.

Notre exposé présente les étapes de notre recherche qui ont mené à la définition d'un modèle. Dans le premier chapitre, nous décrivons le domaine d'application de nos recherches, à savoir un dialogue relatif à des renseignements administratifs, en axant plus spécialement notre attention sur les échanges d'informations entre les différents modules de traitement linguistique au sein d'une architecture de système de dialogue. Dans ce cadre, nous définissons un gestionnaire pour les hypothèses lexicales qui résultent d'une phase descendante d'analyse. Les problèmes posés par une extension possible de ce gestionnaire aux autres types d'information conduisent à critiquer certains aspects de l'architecture envisagée.

Dans le deuxième chapitre, nous faisons un inventaire des modèles existants dans différents domaines des sciences cognitives, qui pourraient servir de références pour notre approche. Nous sommes conduits à refuser certaines voies, alors qu'apparaissent des lignes de recherche privilégiées pour une nouvelle vision d'un système en situation de dialogue.

Cette progression nous amène à choisir un axe d'étude qui touche tous les niveaux de traitement d'un dialogue : le temps. Dans le troisième chapitre, nous proposons un modèle simplifié du temps par rapport aux théories existantes et nous exposons en détail les principaux objets et relations que nous introduisons. Nous confrontons alors ce modèle avec les différents domaines de traitement, à la fois linguistiques et de raisonnement, pour montrer qu'il est possible d'observer tous ces phénomènes dans le même cadre de travail que procure une étude temporelle.

Dans une dernière partie (chapitre 4), nous discutons les principaux problèmes d'implémentation relatifs à ce modèle en faisant ressortir les différentes stratégies qui peuvent être mises en œuvre en situation réelle.

1 Les échanges dans un systeme de Dialogue Oral Homme-Machine.

1.1 Aperçu des difficultés.

Parler de dialogue homme-machine, même en laissant de côté les problèmes propres au langage parlé, semble une gageure quand on essaie d'appréhender les entités impliquées dans cette activité. Il s'agit en fait de faire communiquer deux systèmes qui ne sont pas vraiment conçus pour cela :

- une machine, c'est à dire un ordinateur, dont les caractéristiques principales sont de reposer sur une structure et un fonctionnement extrêmement simples, et qui se trouve être utilisé à l'heure actuelle essentiellement pour des tâches répétitives dont son concepteur, l'homme, maîtrise parfaitement les mécanismes ou algorithmes.
- un être humain, dont la structure est extrêmement complexe, en particulier en ce qui concerne ses capacités cognitives que de nombreuses sciences, comme nous le verrons, s'efforcent de comprendre sous des angles très divers.

Enfin il faut tenir compte d'un environnement qui comprend les deux entités précédemment citées et dont les formes par lesquelles il peut être perçu sont très variées. Les sons, images, odeurs, pressions sont autant de supports qui peuvent être appréhendés à différentes échelles par l'homme et la machine. En l'état actuel des choses, les capteurs, ainsi que les programmes de reconnaissance de ces signaux ne sont pas assez développés pour que la machine dispose de la quantité d'information accessible à l'homme pour ses activités cognitives.

Dans ces conditions, pour définir la notion de dialogue, il est indispensable de considérer des situations de communication plus naturelles, et ici, la référence consiste à étudier ce qui se passe entre deux êtres humains. Ce dernier point de vue est d'autant plus important, que l'on espère réaliser des machines pouvant dialoguer avec des humains et non l'inverse, ceci afin d'éviter une période d'apprentissage trop importante à un utilisateur, comme cela se passe pour les usagers actuels de l'informatique (pour ne pas parler des informaticiens eux-mêmes). Il reste néanmoins une interrogation à laquelle nous ne prétendons pas répondre : le modèle homme-homme est il un bon modèle pour la communication homme-machine ([Amalberti 88]) ?

En première approximation, on peut limiter un dialogue entre deux humains au seul "canal verbal" (cf [Pierrel 81] pp.23-25). Que ce soit par l'oral ou par l'écrit, dans le

cadre de systèmes électroniques tels que le minitel, de nombreuses informations peuvent être échangées par le simple usage d'une langue commune aux deux intervenants. Cependant, il est bien rare que l'on assiste à des situations de dialogue aussi simplifiées. En effet, il suffit de mettre deux personnes l'une en face de l'autre pour observer que de nombreux signes extra-linguistiques, tels que regards, mimiques ou bruits divers, sont utilisés pour communiquer. Ces signes apparaissent d'autant plus que le contenu du dialogue fait lui-même référence à l'environnement des deux interlocuteurs. L'information effectivement transmise par le langage peut alors se révéler extrêmement réduite, et parfois limitée presque uniquement à des références à cet environnement, qui sont alors accompagnées de gestes évocateurs, comme dans la scène que l'on peut s'imaginer à partir du dialogue suivant :

Dialogue 1.1.

- Eh bien? demanda son compagnon.
- C'est là.
- C'est là ! qu'est ce que vous me chantez ?
- Oui là, devant nous.
- Devant nous ! Dites donc, Danègre, il ne faudrait pas...
- Je vous répète qu'elle est là.
- Où ?
- Entre deux pavés.
- Lesquels ?
- Cherchez.
- Lesquels ? répéta Grimaudan.
Victor ne répondit pas.
- Ah ! parfait, tu veux me faire poser, mon bonhomme.
- Non... mais... je vais crever de misère.
- Et alors, tu hésites ? Allons, je serai bon prince. Combien te faut-il ?
- De quoi prendre un billet d'entrepont pour l'Amérique.
- Convenu.
- Et un billet de cent francs pour les premiers frais.
- Tu en auras deux. Parle.
- Comptez les pavés, à droite de l'égout. C'est entre le douzième et le treizième.
- Dans le ruisseau ?
- Oui, en bas du trottoir.

Maurice Leblanc, La perle Noire.

Enfin, même en se limitant à la seule communication linguistique, la complexité d'un dialogue peut provenir du nombre de sujets abordés, nécessitant de la part des deux interlocuteurs une quantité importante de connaissances pour reconnaître, interpréter, comprendre chacun des énoncés de l'autre. L'homme fait preuve à cet égard de capacités particulièrement importantes pour gérer toute ces connaissances et les réutiliser au moment adéquat dans un dialogue, et de manière plus générale, pour réagir et s'adapter à l'univers qui l'entoure.

Par rapport à ces importantes facultés, les perspectives qu'offre l'outil informatique sont bien minces : nous disposons de peu de mémoire, de moyens de calcul extrêmement lents et ne sachant effectuer quasiment que des opérations sur des nombres. La mise en œuvre d'un système de dialogue entre une machine et un être humain va donc nécessiter une simplification de tous les mécanismes mis en cause, particulièrement en ce qui concerne l'étendue du domaine des dialogues envisagés. Par ce biais, nous espérons réduire la quantité de connaissances qu'il faudra manipuler mais aussi supprimer certaines difficultés qui peuvent être dues, par exemple, à la taille du langage utilisé.

Ce dernier point ne résout pas obligatoirement tous les problèmes. En effet, certaines particularités de la langue, que l'on nomme parfois exceptions, sont souvent l'expression de mécanismes sous-jacents au niveau cognitif qui permettraient éventuellement d'expliquer des phénomènes plus courants qui apparaissent dans les mécanismes de compréhension d'une langue. Réduire de manière trop importante la complexité du langage utilisé risque alors d'interdire l'extension du système que l'on aurait ainsi construit.

Le cadre de l'ordre des mots dans la phrase est représentatif à cet égard. En effet, nous ne sommes pas en mesure d'établir un ensemble de règles fini qui représente toutes les variantes que chaque écrivain ou locuteur introduit, surtout en situation de dialogue où seuls comptent les groupements apportant du *sens* au discours (Dialogue 1.1 : "Devant nous", "Entre deux pavés", "Et un billet...").

1.1.1 Description du domaine.

Le choix du domaine d'application de nos recherches s'est effectué dans notre équipe en fonction de différents critères. Comme nous l'avons signalé, l'univers du discours doit être réduit par rapport à une conversation courante entre deux personnes quelconques. Cependant, il doit rester suffisamment vaste pour que de nouvelles idées de recherche puissent voir le jour face à des difficultés qui n'avaient pas été nécessairement traitées jusqu'ici. Le choix s'est donc porté sur la simulation d'un centre de renseignements administratifs relatifs à des informations proches de celles rencontrées dans les "pages roses de l'annuaire téléphonique" français. L'utilisateur potentiel d'un tel système est en situation de demandeur, le système doit dialoguer avec lui pour préciser l'objet exact de sa question et fournir alors une réponse personnalisée qui le satisfera. Afin d'illustrer la

forme que peuvent prendre dans ce contexte les échanges linguistiques entre l'utilisateur et la machine, nous pouvons observer un exemple de dialogue :

Dialogue 1.2.

(S : système, L : locuteur)
S1 : Centre de renseignement administratif, Bonjour!
L1 : Bonjour, A qui dois-je m'adresser pour une carte d'identité ?
S2 : A un poste de police ou à la mairie.
L2 : Quelles sont les formalités ?
S3 : Etes-vous mineur ?
L3 : Non,non.
S4 : Vous devez présenter votre livret de famille ou une fiche d'état civil et fournir un timbre fiscal.
L4 : Combien coûte un timbre fiscal ?
S5 : 115 F.
L5 : Merci.
S6 : Désirez-vous autre chose ?
L6 : Non. Merci. Au revoir.
S7 : Au revoir.

Plusieurs faits peuvent être observés à partir de ce dialogue. Tout d'abord on mesure la différence avec les systèmes d'information couramment rencontrés, dans le sens où la machine n'a pas l'exclusivité des initiatives de dialogue, ce qui ne laisserait alors à l'usager que le choix entre plusieurs réponses prédéfinies. Il ne s'agit pas de situations où le système n'attend de l'utilisateur que des valeurs pour des paramètres particuliers et où le nombre possible de besoins de l'utilisateur est suffisamment réduit pour pouvoir lui proposer un système de menus ou une suite de questions fermées ; etc. Dans le cas qui nous intéresse, il est effectivement possible de parler de dialogue naturel entre l'homme et la machine : chacun expose ses besoins à l'autre pour faire converger la discussion vers une satisfaction réciproque. Ainsi, dès le premier énoncé du locuteur (L1), le système doit s'adapter à une requête particulière, mais peut toujours, si le besoin s'en fait sentir, reprendre le contrôle du dialogue (S3) et effectuer une demande précise à l'usager. Le rôle d'un tel système peut être assimilé à celui d'un standardiste humain pouvant s'adapter avec plus ou moins de succès aux circonstances d'un dialogue donné. Cependant, on ne se posera pas les éventuels problèmes économiques et donc humains qu'un tel remplacement pourrait engendrer.

Dans un deuxième temps, on peut évaluer la complexité de la langue utilisée en constatant la simplicité des structures de phrase de ce dialogue. Ce sont essentiellement des phrases simples, déclaratives ou interrogatives, accompagnées éventuellement de groupes prépositionnels (L1 : "pour une carte d'identité") qui précisent le contexte local de l'énoncé. Au sens des grammaires formelles, le niveau de récursion est principalement de 1, parfois de 2, mais il est très rare que l'on ait à faire à des enchâssements répétés de

syntagmes comme dans : "le chat du frère de l'ami de ma voisine". La difficulté de compréhension de tels dialogues ne semble pas résider dans la complexité syntaxique des énoncés, mais plutôt dans leur diversité d'interprétation en fonction d'un contexte qui garde constamment une grande importance. Nous pouvons constater que la compréhension d'un énoncé particulièrement simple tel que L2 : "Quelles sont les formalités ?" nécessite le déclenchement de mécanismes d'inférence qui mettent celui-ci en rapport avec la demande précédente relative à une carte d'identité .

Au vu de ces premières remarques, apparaît une notion très importante pour la mise en œuvre d'un système de dialogue homme-machine, celle de sous-langage ([Pierrel 87] p.65). Elle permet de décrire, dans l'étendue d'un langage naturel quelconque, les restrictions lexicales, syntaxiques ou sémantiques qui apparaissent dans un contexte finalisé particulier, correspondant à la réalisation d'une tâche spécifique dans un domaine d'application bien défini. Les sous-langages se caractérisent essentiellement par un vocabulaire non-grammatical assez réduit et pour lequel chaque entrée particulière est dans la plupart des cas monosémique et permet donc de retrouver très rapidement un concept de l'univers du dialogue. Toutes ces restrictions peuvent être obtenues à partir de l'étude d'un corpus représentatif du domaine de la tâche, ce qui différencie d'ailleurs un sous-langage par rapport à un langage naturel. Diverses études ont été réalisées au sein de notre équipe pour délimiter le sous-langage particulier, propre au domaine que nous nous étions fixé ([Roussanaly 86], [Deville 87b]). Une analyse poussée a été réalisée par des linguistes belges dégageant plus particulièrement les caractéristiques sémantiques des formes prédicatives pouvant apparaître dans une demande de renseignements administratifs ([Deville 86], [Deville 87a]). Nous reviendrons sur ce point lors de l'étude des modèles linguistiques utilisés dans le système de dialogue oral homme-machine servant de base à nos recherches.

Enfin, nous pouvons situer l'ambition du projet présenté en remarquant qu'il concerne un domaine relativement vaste par rapport aux systèmes de dialogue proposés jusqu'ici. Sans parler d'ancêtres lointains tels que SHRDLU [Winograd 73], les systèmes actuels développés tant au LIMSI à Orsay ([Aouati 84],[Bomerand 88]) qu'au CNET à Lannion ([Siroux 85]) concernent des applications du même type (demande de renseignements), mais correspondent à des situations où les possibilités de demandes sont plus réduites comme dans les systèmes ESOPE au LIMSI ou KEAL au CNET qui concernent des applications du type standard téléphonique. Cependant, ces systèmes présentent l'intérêt d'être arrivés à un stade pré-industriel, puisque, comme le système de dialogue pilote-radar du LIMSI, ils sont déjà testés en situation réelle.

1.1.2 Problèmes spécifiques dans un dialogue oral homme machine.

En fonction de ce qui vient d'être dit, nous nous trouvons confrontés à toute une série de problèmes liés à différents phénomènes acoustiques, linguistiques ou cognitifs qu'il est nécessaire de résoudre pour proposer réellement un dialogue naturel à un utilisateur.

Dans le cadre d'un dialogue homme-machine quelconque, un système de reconnaissance doit recueillir les énoncés successifs de l'utilisateur, les analyser en fonctions de ses connaissances linguistiques et enfin, construire une représentation interne de cet énoncé utilisable pour effectuer ultérieurement des inférences sur l'univers de la tâche. C'est à cette dernière condition, qu'il sera possible d'affirmer que le système a effectivement *compris* le message qui lui a été transmis. L'activité qui vient d'être décrite pourrait être assimilée à une simple reconnaissance de phrases, si chaque énoncé n'était en dépendance étroite avec ceux qui l'ont précédé, que ce soient les autres interventions du locuteur ou celles de la machine, puisque, par définition, chaque intervenant d'un dialogue tient compte de ce que l'autre lui a dit. Dans le genre de dialogue qui nous intéresse, les liens les plus typiques qui peuvent être observés sont des couples question-réponse, où le deuxième énoncé n'a de sens qu'en fonction de celui qui le précède. A titre d'illustration, dans le dialogue précédent nous avons un exemple standard de question ouverte, où la réponse ne veut strictement rien dire sans référence à la question posée.

S3 : Etes-vous mineur ?
L3 : Non,non.

Dans un dialogue quelconque (cf. dialogue 1.1), les liens entre les énoncés peuvent ne pas se limiter à des couples adjacents, mais la compréhension d'un élément peut nécessiter le rappel d'un contexte beaucoup plus ancien; par exemple :

- Et un billet de cent francs pour les premiers frais.
- Tu en auras deux. Parle.

Ici, alors que la première partie de l'énoncé est directement interprétable à partir de ce qui le précède, la deuxième : "Parle", doit, pour être comprise, être restituée dans le contexte des tout premiers échanges du dialogue, bien avant les tractations concernant le prix à payer pour obtenir l'information demandée.

Bien que nous nous soyons situés dans un cadre plus restreint, tout dialogue forme un ensemble dans lequel s'insère chaque énoncé pour pouvoir être interprété. En particulier, lorsqu'un thème général du dialogue ressort - dans le dialogue 1.2, il s'agissait de carte d'identité - celui-ci se maintient tant qu'un nouveau thème n'apparaît pas, ou il peut être modifié, précisé, suivant le degré d'avancement du dialogue. Par exemple si un dialogue porte sur la demande d'un papier d'identité pour partir à l'étranger, en fonction du pays d'accueil, l'échange va plus précisément s'orienter vers les problèmes de carte d'identité ou sur les formalités qu'il faut effectuer pour faire une demande de passeport.

Il apparaît donc nécessaire de gérer ce contexte de dialogue de façon continue, ce qui implique deux opérations, à savoir la construction de ce contexte (que l'on nommera historique du dialogue) et la restitution de l'information pertinente pour la compréhension d'un énoncé précis. De façon générale, quand le dialogue fait intervenir deux humains cette dernière opération s'effectue avant même que le nouvel énoncé soit apparu. Chaque intervenant est ainsi en mesure d'effectuer des *prédictions* sur les énoncés à venir, réduisant ainsi l'espace de recherche potentiel dans lequel ceux-ci vont s'insérer. Ce phénomène représente une activité importante, puisqu'elle permet de pallier certaines difficultés spécifiques à l'usage de la parole dans un dialogue.

En effet, la parole introduit de nombreux problèmes par rapport à un simple dialogue écrit. Comme tout autre phénomène physique, elle est continue, contrairement aux éléments linguistiques manipulés qui, eux, sont séparés en éléments discrets. De plus, un message donné n'est pas exprimé de manière unique par tous les locuteurs. Chacun a sa propre façon de parler en fonction de son âge, de son sexe ou de sa provenance géographique qui introduit un accent souvent très prononcé. Cette variabilité importante se retrouve au niveau d'un même locuteur dont la voix est sujette à diverses modifications en fonction de la fatigue ou d'une quelconque maladie. La difficulté consiste donc à reconnaître, au sein d'un signal extrêmement variable et bruité, des éléments linguistiques précis.

Une autre différence entre un dialogue écrit et un dialogue oral réside dans une particularisation du langage lui-même. Alors que le langage écrit se traduit par des phrases relativement longues et d'une certaine manière, grammaticalement complètes, le langage parlé est fortement simplifié, comme nous pouvons l'observer dans les exemples de dialogue 1.1 et 1.2 (bien que le deuxième, exemple d'école, soit proche d'un dialogue écrit). Les phrases sont souvent beaucoup plus courtes, très élidées au niveau phonétique et surtout elles s'avèrent incomplètes, puisque seules les informations réellement utiles au

dialogue sont exprimées. Cela pose certains problèmes d'interprétation et les ambiguïtés éventuelles qui apparaissent doivent être résolues par des procédures de dialogue spécifiques qui peuvent conduire à une interrogation du locuteur.

Alors que les difficultés précédentes concernaient essentiellement le locuteur et sa manière de produire des énoncés, l'environnement d'énonciation intervient lui aussi en introduisant dans le signal reçu des bruits divers rendant d'autant plus difficile la compréhension. Face à un énoncé inaudible, parfois partiellement reconnu, il est là encore nécessaire de relancer le locuteur, pour qu'il répète ce qu'il vient de dire, ou préférablement, pour qu'il reformule sa phrase afin d'obtenir une redondance sémantique qui permet de mieux cerner le contenu de son énoncé initial.

1.1.3 Connaissances à mettre en œuvre.

Les différents problèmes présentés succinctement au paragraphe précédent semblent d'emblée concerner des domaines divers de connaissances. De plus, nous avons souvent parlé de reconnaissance et de compréhension, sous-entendant par là une possible sériation de ces phénomènes. Il semble évident que, pour traiter autant de difficultés dans le cadre d'un dialogue, il soit nécessaire de faire intervenir toutes les connaissances qui sont à notre disposition, ceci sans distinction d'origine, en empruntant des résultats qui proviennent d'autres sciences que les sciences informatiques, en particulier, en allant voir du côté des sciences humaines, telles que la psychologie, la linguistique, etc...

Nous avons observé l'importante incertitude qu'il y avait à interpréter un énoncé à cause des aspects propres à la parole. Dans ce cadre, notre but est donc double. Premièrement, pour chaque ambiguïté apparaissant dans une phase de reconnaissance, l'utilisation de connaissances adéquates va permettre la réduction du nombre d'hypothèses possibles et donc la simplification d'une compréhension ultérieure. Deuxièmement, nous avons fait remarquer qu'il serait intéressant d'effectuer des prédictions sur les énoncés à venir en fonction des contraintes qu'introduisent les connaissances disponibles. Ces deux points rendent notre manière d'aborder ces problèmes plus proche des préoccupations de l'intelligence artificielle que de celles de techniques tournant autour de la reconnaissance des formes. Loin de vouloir faire un automate de reconnaissance, les opérations de discernement et de prédiction nous rapprochent plutôt de la définition d'un schéma cognitif de compréhension.

Enfin, puisque le système doit raisonner sur les objets de l'univers de la tâche, nous ne pouvons nous contenter d'une représentation simpliste du contenu de chaque énoncé (une seule structure syntaxique par exemple). Au contraire, nous devons construire, à partir du signal de parole, un modèle de cet univers d'un haut niveau d'abstraction, en intégrant toutes les informations linguistiques (ou autres) intermédiaires.

Parler de niveaux donne déjà une indication sur la manière dont on peut cerner ces connaissances. Puisqu'il semble impossible, à l'heure actuelle, de percevoir globalement l'ensemble des informations utilisables pour interpréter un énoncé dans un dialogue, une solution envisageable est d'adopter une vision modulaire, afin d'extraire une série de sous-problèmes spécifiques, qui pourront être résolus de manière indépendante. Ceci est d'autant plus intéressant dans un contexte informatique que l'on peut imaginer assigner à chacun de ces modules une unité de calcul particulière.

En marge de ce souci stratégique, séparer les informations linguistiques et cognitives est historiquement encouragé par les études qui ont été réalisées dans différents domaines et qui concernent des aspects spécifiques. Schématiquement, on rencontre des phonéticiens spécialisés dans l'étude des sons de la langue, des linguistes qui analysent les problèmes plus syntaxiques, des sémanticiens ou encore, des psychologues qui analysent comment des sujets se représentent les informations véhiculées dans des phrases. On peut ainsi extraire différentes catégories d'informations, qui, semble-t-il, doivent être toutes intégrées à un système de compréhension de dialogue.

En bas d'une échelle partant du signal de parole jusqu'à une représentation interne, apparaît la *phonétique*. Elle est plus particulièrement chargée de transformer un flux continu d'informations en unités symboliques : les phonèmes. En fait, suivant le type de données que l'on considère comme étant fondamental, il peut s'agir de segments plus larges que les phonèmes, diphones ou syllabes ou au contraire d'éléments de taille inférieure. Dans un système de compréhension, un module phonétique permet donc de passer d'une représentation dont les frontières sont imprécises, à cause des phénomènes de transitions entre sons et de coarticulations, à une représentation incertaine (incertitude quant à la présence de tel phonème sur telle zone du signal) plus utilisable pour les niveaux supérieurs.

Intervient ici une source de connaissances très importante, que nous traiterons dans le détail ultérieurement : le *lexique*. Celui-ci fait correspondre à des phonèmes des mots de la langue. Cependant, on attache souvent de nombreuses informations supplémentaires au

niveau lexical, que ce soient des informations syntaxiques ou sémantiques. Ainsi, M. Gross (77) considère que les contraintes syntaxiques de la langue peuvent être ramenées au niveau du mot, en inventoriant toutes les situations possibles où un item lexical intervient.

Déjà, nous rencontrons les écueils d'une définition strictement modulaire des connaissances linguistiques. La *syntaxe*, dont nous venons de voir le lien possible avec le lexique, a longtemps été au centre de la discussion sur la modularité de la langue. A la suite des travaux de Chomsky et de l'école générative, elle a pendant un moment acquis un statut d'autonomie quasi-total. On pouvait penser alors que de simples règles de réécriture permettraient d'analyser la structure de phrases en terme de syntagmes ou de catégories grammaticales, sans faire appel à d'autres connaissances. Pourtant, il était impossible d'expliquer que, dans les contextes courants, on n'accepte pas :

? la casserole mange la souris

par rapport à :

le chat mange la souris.

La *sémantique* est le domaine d'analyse de ces phénomènes, c'est à dire des contraintes existant entre les éléments d'une phrase, non pas du fait de leur position, mais à cause de leurs significations propres qui doivent rester compatibles. Cependant, la situation n'est pas aussi simple que pourrait le laisser penser une telle définition. En effet, l'inacceptabilité de telle ou telle phrase au niveau sémantique est souvent beaucoup plus complexe à établir qu'au niveau syntaxique. Il n'est pas rare que dans des exemples aux limites du possible, cette détermination conduise à des débats passionnés. Ainsi, pour le seul exemple précédent, pourquoi ne pas imaginer une situation de discours particulière - un monde à la Lewis Caroll conviendrait parfaitement - où une casserole peut manger une souris. Selon l'expression de Guy Deville, parler en ces termes serait commettre un "viol pragmatique" sur la sémantique. Pourtant, il ne semble pas que l'on puisse effectivement donner le sens d'une phrase sans vraiment établir un environnement de discours qui lui corresponde.

La *pragmatique* justement, dernier module de l'échelle, pose plus de problèmes existentiels. Afin de la différencier de la sémantique, elle a été définie parfois comme la science du contexte, c'est-à-dire celle qui précise comment une phrase peut se transformer en un acte réel d'énonciation. Dans notre perspective, la pragmatique consiste plus précisément à insérer un énoncé dans un contexte de dialogue et donc à faire le lien avec les objets du discours manipulés par la partie raisonnement du système.

Une fois définis, ces modules ne présentent un intérêt pour nous que s'il existe, pour chacun d'eux, des modèles implémentables sur une machine. De plus, l'hypothèse de modularité nous impose de définir un protocole d'interaction entre ces différents processus afin de former un système cohérent, réalisant convenablement le but que nous nous sommes fixé, à savoir : établir un dialogue naturel entre l'homme et la machine. Nous allons maintenant exposer une proposition d'un tel agencement menant à la définition d'une architecture de système, puis préciser comment, dans ce cadre, nous pouvons étudier plus particulièrement les échanges portant sur les informations lexicales.

1.2 Définition d'une architecture.

Avant de décrire dans le détail une architecture possible pour un système de dialogue oral homme-machine, rappelons que notre objectif principal est d'obtenir une organisation qui fasse effectivement converger la reconnaissance d'un énoncé vers une solution satisfaisante en utilisant toutes les connaissances possibles à différents niveaux, du signal jusqu'à une représentation interne relativement abstraite. Dans un premier temps nous présenterons le paradigme général de l'organisation de ce système, pour ensuite décrire dans le détail les différents modules qui interviennent, ainsi que les échanges entre ceux-ci, dont nous verrons qu'ils ont beaucoup d'importance dans notre approche.

1.2.1 Principes généraux de l'organisation du système.

A la suite de l'expérience acquise par les membres de l'équipe "parole" du CRIN dans le cadre successif des projets de reconnaissance Myrtille I et Myrtille II développés par J.M.Pierrel (83), une nouvelle architecture a progressivement vu le jour qui, comme nous le verrons, diffère profondément des systèmes précédents. En fonction des connaissances que nous avons considérées comme nécessaires pour analyser correctement un énoncé dans un dialogue, nous avons adopté une hypothèse de travail consistant à assimiler le système de dialogue à un ensemble de modules de reconnaissance possédant une certaine autonomie. Chaque module est vu comme un processus indépendant fonctionnant suivant un modèle de type producteur-consommateur. Schématiquement, en fonction de ses propres entrées, un module va produire un certain nombre de résultats sur ses sorties en utilisant ses facultés de décision qui peuvent être résumées ainsi :

- chaque module possède sa base de connaissance, sans avoir accès à celle des autres, ce qui permet une gestion répartie de l'information.

- en fonction de ces connaissances et de l'information disponible en entrée d'un module, celui-ci va pouvoir construire un ensemble particulier de représentations d'un énoncé. Il ne voit donc l'énoncé que sous une forme donnée, et ignore de ce fait les représentations adoptées par les autre modules.

- enfin, chaque module possède sa propre stratégie (ascendante, descendante ou mixte) et peut la modifier à loisir en choisissant vers quels modules il envoie le résultat de ses analyses.

A partir de ces quelques remarques, il est possible de situer l'architecture envisagée par rapport aux systèmes développés jusqu'ici dans d'autres laboratoires. Cette comparaison est relativement aisée puisque la plupart de ces systèmes ont adopté eux aussi une approche modulaire, mais structurée suivant des principes un peu différents. Une exception à cette règle pourrait être le système Harpy (Lowerre, 80) qui, après une étape de compilation de ses connaissances, intègre en un seul niveau tous les modules de reconnaissance. Nous signalons que la notion d'intégration qui apparaît ici est différente de celle que nous introduirons lors de notre discussion sur de nouvelles architectures, où l'homogénéisation des connaissances que nous demanderons ne sera pas une simple optimisation informatique. Le système Harpy mis à part, nous pouvons donc distinguer deux grandes classes de systèmes de reconnaissance et de compréhension de la parole (et parfois systèmes de dialogue) :

- dans les systèmes hiérarchiques, la stratégie est imposée par les choix relatifs à l'architecture d'ensemble. Cette stratégie dépend alors d'un superviseur décidant à tous les niveaux des choix de reconnaissance à effectuer - Hwim (Wolf, 1980) en est un exemple typique - ou tout simplement, de l'organisation des différents modules entre eux, ce qui impose une stratégie fixe et définie une fois pour toute. Dans ce dernier cas, nous pouvons situer les systèmes à stratégie ascendante tels que Keal [Siroux 85], ou le système de dialogue pilote/avion réalisé au LIMSI [Bornerand 88], et ceux à stratégie descendante (guidés par la syntaxe par exemple) pour des systèmes plus anciens tels que ESOPE [Mariani 78] ou à Nancy, Myrtille I ([Pierrel 78]).

- à l'opposé, on observe des systèmes non-hiérarchiques, dont l'exemple le plus connu est la structure en *blackboard* du système Hearsay II ([Erman 80b]). Dans ce dernier cas, le but de ses concepteurs était de créer un modèle assez souple permettant de mettre au point diverses organisations et en particulier, de développer tout un éventail de stratégies afin d'évaluer l'intérêt de chacune d'entre elles. Ces stratégies sont malgré tout décidées en dehors de chaque module par un mécanisme particulier, ce qui est une limite pour une réelle autonomie. Dans ce cadre d'analyse, notre architecture peut être considérée comme non hiérarchique, mais elle diffère des structures de type blackboard par la grande indépendance réservée à chaque module de reconnaissance.

1.2.2 Description de l'architecture et modèles linguistiques utilisés.

La figure 1.1 nous montre l'organisation du système de dialogue qui s'articule autour de cinq modules principaux, et de deux pôles d'échanges que nous justifierons ultérieurement. Chaque module vérifie les conditions que nous évoquions, c'est à dire autonomie de gestion et de traitement des informations. De plus, certains d'entre eux reposent sur des modèles linguistiques spécifiques que nous détaillerons dans la suite.

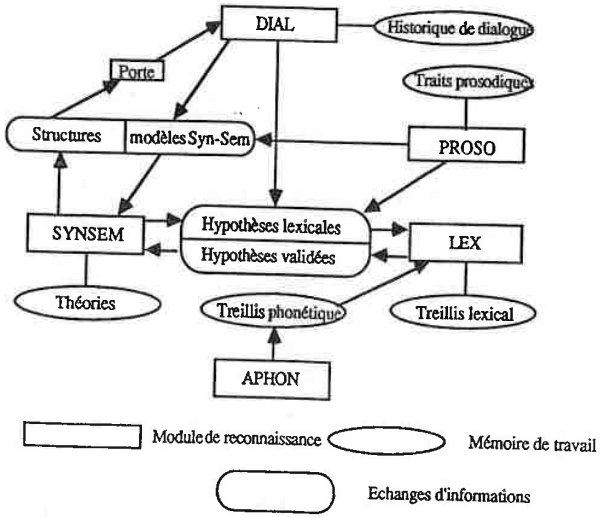


Figure 1.1 : une architecture de système de DOHM.

Partie fondamentale d'un système de reconnaissance, le module de décodage acoustico-phonétique (APHON) repose sur un ensemble de recherches effectuées au CRIN autour de deux axes principaux, SNORRI [Laprie 88], un système de traitement et d'analyse de la parole, et APHODEX [Fohr 87], un système expert de décodage phonétique. Dans le cadre de notre système, les traitements acoustico-phonétiques sont purement ascendants et peuvent être assimilés à l'effet d'une seule unité de transformation du signal jusqu'à un treillis de phonèmes. A l'heure actuelle, les résultats obtenus en décodage acoustico-phonétique demeurent trop faibles pour que l'on puisse être certain des éléments linguistiques manipulés ultérieurement. Ceci justifie en particulier l'importance que nous accordons aux prédictions pouvant être effectuées par le système.

Nous pouvons au passage citer le module de détection d'indices prosodiques, PROSO, et les éventuels résultats qu'il pourrait produire dans un avenir plus ou moins lointain. Les principaux éléments que nous possédons à l'heure actuelle semblent en effet limiter l'usage de la prosodie à la détection de frontières de syntagmes ou de mots et donc à la production d'hypothèses temporelles pour l'analyse lexicale ([Carbonell 88]). Bien plus, le lien que nous avons indiqué entre PROSO et les analyseurs syntaxico-sémantiques est nettement hypothétique, puisqu'il s'agirait de prédire certaines structures énonciatives (question, affirmation, etc...) en fonction par exemple des variations de la fréquence fondamentale, extraite directement du signal. Malgré tout, l'obtention de seules frontières de mots est un aspect particulièrement important, puisque cela permettrait de réduire de façon exponentielle la taille des treillis lexicaux produits de façon ascendante, par exemple.

LEX, le module d'analyse lexicale, réalise plusieurs fonctions suivant la nature des informations qui sont disponibles sur ses entrées, conformément au schéma que nous avons décrit. Dans le cas d'un processus ascendant, ce module va ainsi transformer un treillis phonétique en provenance de APHON en un treillis lexical, ou plutôt en un ensemble de mots candidats qu'il va pouvoir proposer directement à SYNSEM. Cette fonction est réalisée grâce à des macro-classes de phonèmes qui permettent d'accéder rapidement à une base de données lexicales structurée suivant ce principe. A partir des seules classes 'Plosives', 'Fricatives' et 'Vocaliques' on peut ainsi réduire d'un facteur 20 à 50 le nombre de candidats possibles sur une portion de signal ([Romary 88]). Inversement, des hypothèses lexicales en provenance des autres modules peuvent nécessiter une vérification descendante sur le treillis phonétique. Cette opération est facilement effectuée par programmation dynamique sur les chaînes de phonèmes candidates, afin de donner, éventuellement, le meilleur positionnement temporel de ces chaînes. Comme nous pouvons le constater, il ne s'agit pas ici d'utiliser des modèles particuliers de description du lexique, mais de définir des traitements spécifiques caractérisés par leur efficacité. Ce même module a d'ailleurs servi sous une forme similaire pour d'autres systèmes de dialogue en développement au CRIN, à savoir DIAPASON, un système de dialogue avec une console sonar, dont une première version est présentée dans [Alinat 87] et PARTNER [Morin 87], une interface entre un utilisateur et un système de messagerie électronique. Enfin, l'origine des différentes hypothèses lexicales arrivant à LEX pouvant être très diverse, nous verrons qu'il est nécessaire de les gérer afin de les ramener à un format commun directement exploitable pour une vérification sur un treillis de phonèmes.

Abordons maintenant les modules relatifs à des informations complexes et donc de prime abord plus structurées, SYNSEM et DIAL.

SYNSEM cumule en un seul module les fonctionnalités d'analyse syntaxique et sémantique. Il regroupe donc l'ensemble des connaissances structurelles relatives à la langue utilisée. La raison ayant conduit à la définition d'un seul module est le lien étroit existant entre les structures purement syntaxiques et les éléments manipulés dans des formes prédicatives telles que celles qui apparaissent dans la plupart des modèles sémantiques courants.

Dans le cas présent, le module d'analyse structurale repose principalement sur deux modèles linguistiques et/ou informatiques : des Réseaux à Nœuds Procéduraux (RNP) pour la syntaxe et une grammaire de cas pour la sémantique.

Les RNPs ont été développés par J.M. Pierrel dans le cadre du projet Myrtille II [Pierrel 81] afin d'étendre le concept habituel de grammaire , pour aboutir à un ensemble de réseaux d'analyse. Pour chacun des nœuds de ces réseaux, il est possible de déclencher des procédures spécifiques permettant de vérifier en particulier certains indices pertinents directement sur le signal. Pour SYNSEM, un compilateur de tels réseaux a été développé de manière à pouvoir paramétrer la partie syntaxique du système en fonction d'une étude de corpus par exemple. Pour chaque énoncé, les analyseurs gèrent donc un ensemble de THÉORIES, c'est à dire des analyses partielles fournies par les RNPs, dont la gestion, suivant des stratégies spécifiques, permet de faire, soit des hypothèses descendantes (lexicales), soit des hypothèses ascendantes sous la forme de structures candidates pour le module de dialogue. Bien évidemment, ces théories ne contiennent pas uniquement des informations syntaxiques relatives à l'énoncé en cours d'analyse, mais aussi une représentation sémantique associée, qui est construite en parallèle.

Le modèle sémantique que nous utilisons - une grammaire de cas - est une variante du modèle proposé par Fillmore en particulier (68). La présente adaptation fait suite à une étude effectuée par Guy Deville et Hans Paulusen ([Deville 86], [Deville 87a]) et poursuivie depuis dans le cadre de la thèse de Guy Deville (89). Cette analyse a conduit à la définition d'un ensemble de prédicats pertinents pour notre application (renseignements administratifs) ainsi que les cas correspondants. Plutôt que de définir un prédicat pour chaque verbe, nom ou adjectif du lexique, la grammaire de cas s'appuie sur un certain nombre de primitives sémantiques, correspondant aux grandes catégories de situations

pouvant apparaître dans une application donnée. A titre d'exemple, au verbe "Signer" peut être associée la primitive 'ACTOBJET' dont les cas principaux sont l'agent (AGT), l'objet (OBJ) et, de manière optionnelle, un Bénéficiaire (BEN). Une autre manière de définir une primitive sémantique repose sur la description des contraintes exprimées par un ensemble de traits. Ainsi 'ACTOBJET' est une action contrôlée (+CTR), dynamique (+DIN), faisant intervenir au moins une entité animée (+ANI) et caractérisée par une absence de Mouvement (-MVT). A ces structures de traits et de cas est associé un ensemble de règles permettant de passer d'une représentation syntaxique à une représentation équivalente sous forme prédicative.

Le dernier, et probablement le plus important des modules de cette architecture réunit en un seul élément un ensemble de fonctions diverses mais étroitement associées, c'est DIAL. Défini de manière complète dans la thèse d'Azim Roussanaly (88), nous n'allons en décrire que les grandes lignes, ainsi que les éléments pertinents pour cet exposé. Schématiquement, DIAL se charge de traiter toutes les informations relatives au contexte de dialogue :

- il gère l'historique d'un dialogue, c'est à dire les différents échanges qui sont intervenus, ainsi que les éléments fournis par le locuteur au cours de ceux-ci, sous la forme de données brutes (âge, ville etc...), ou de requête (objet de la demande).

- en fonction des informations contenues dans l'historique, DIAL est en mesure d'interpréter les énoncés représentés sous forme prédicative en provenance de SYNSEM. L'interprétation d'un énoncé consiste à retrouver en mémoire les objets désignés dans le discours (références multiples à un objet) mais aussi résoudre les références anaphoriques. L'exemple le plus courant consiste à reconnaître derrière 'je', le locuteur lui-même, mais il est aussi nécessaire de savoir que l'on parle d'un même élément au cours d'un dialogue, quand on observe successivement les expressions 'carte d'identité', 'carte', 'elle' ('...est périmée') etc...

D'autres références à déterminer lors de l'interprétation d'un énoncé sont les ellipses, c'est à dire des portions de phrase qui ne sont pas répétées d'un énoncé à l'autre et qu'il est nécessaire de reconstituer pour pouvoir intégrer l'énoncé à l'espace de représentation. Ainsi, dans :

L4 : Combien coûte un timbre fiscal ?
S5 : 115 F.

le prédicat 'coûter' a disparu, bien que sous entendu par le contexte. L'historique du dialogue fournit alors directement le contexte prédicatif dont le groupement "115 F" peut être l'un des cas admissibles.

- une fonction particulière allouée à DIAL concerne la gestion de la reconnaissance d'un énoncé particulier. Cette activité est constituée de deux aspects complémentaires. Tout d'abord, puisqu'il est le seul module à posséder une vision globale de chaque dialogue, DIAL peut déclencher l'acquisition d'un nouvel énoncé du locuteur : par exemple juste après une intervention du système. De plus, il doit prendre l'initiative d'arrêter l'analyse de l'énoncé courant, soit quand il estime qu'une interprétation correcte en a été obtenue soit quand les ressources en temps allouées à la reconnaissance ont été épuisées, quelle que soit la qualité de l'analyse en cours. On peut éviter ainsi une trop longue attente du locuteur entre deux énoncés.

- à partir de l'historique du dialogue et de l'énoncé courant, DIAL doit réaliser la tâche principale pour laquelle il est destiné, à savoir satisfaire la requête du locuteur. Pour cela il lui faut tout d'abord raisonner sur les éléments d'informations dont il dispose pour déterminer précisément cette requête et quand celle-ci semble claire, fournir des éléments de réponse en fonction de sa base de çonnaissance.

Indépendamment des opérations purement déductives qui peuvent être effectuées sur sa base de connaissance, DIAL s'appuie sur un modèle particulier de représentation d'un dialogue qui lui permet d'agir convenablement en fonction du type d'informations qui lui manque. Ce modèle s'appuie sur les travaux de l'école Genevoise de linguistique ([Moeschler 85]), et, sommairement, consiste à représenter un dialogue de manière récursive sous forme de phases spécifiques correspondant à la réalisation d'un objectif précis. Au niveau le plus élevé d'un dialogue on peut ainsi distinguer trois grandes phases : une phase d'introduction (salutations et présentations), une phase de clôture, et entre celles-ci, une phase de réalisation de la requête du locuteur. Cette dernière phase peut, en fonction de la complexité de la demande, se décomposer à nouveau en sous-éléments correspondant par exemple à une détermination de la requête, un apport d'informations spécifiques par le locuteur (suite à une demande du système), ou une précision de la réponse du système. Un point particulièrement intéressant dans ce modèle est qu'il permet à chaque niveau du dialogue l'introduction de phases particulières de traitements des erreurs dues à une mauvaise compréhension de la part du système ou du locuteur. Le dialogue entre les deux interlocuteurs n'est donc pas figé, mais autorise à tout instant des ruptures de séquence.

Il est intéressant de constater combien ce modèle de dialogue se rapproche des grammaires existant au niveau de l'énoncé pour structurer les différents éléments qui le composent. Cette ressemblance provient bien évidemment de la similarité des contraintes à exprimer entre des parties de discours qui se succèdent dans le temps tout en se répondant l'un à l'autre. A ce niveau, la séparation de ces traitements similaires en processus autonomes ne semble pas opportune et ce peut être par là même un nouvel argument contre la modularité.

1.2.3 Les échanges dans ce schéma d'architecture.

Nous venons d'étudier une architecture de système de dialogue en fonction des principales unités qui la composent. Nous avons vu que chacune d'entre elles avait besoin, pour fonctionner, d'informations en provenance des autres modules qui doivent donc transiter suivant un certain protocole de sorte que le système ait un comportement d'ensemble convenable. En particulier, nous pouvons remarquer que les informations susceptibles de passer d'un module à un autre sont de type très divers : frontières de mots données par PROSO, mots fournis par LEX, structures syntaxico-sémantiques construites par SYNSEM etc... De plus les informations peuvent être scorées, et il faut donc pouvoir comparer l'influence relative de ces différents scores sur les connaissances d'un module donné.

Si on observe le schéma de l'architecture présenté dans la figure 1.1, deux modules semblent jouer un rôle central par rapport à l'ensemble du système : LEX et SYNSEM. En effet, alors que APHON et PROSO fonctionnent essentiellement de façon ascendante, et que DIAL manipule surtout des informations "conceptuelles", ces deux modules voient transiter autour d'eux beaucoup d'informations que l'on peut scinder en deux grands flux:

- les informations lexicales. Elles peuvent provenir de DIAL, SYNSEM ou PROSO en fonction de contextes d'analyse particuliers, ou de LEX en fonction des phonèmes trouvés sur le signal.

- les informations syntaxico-sémantiques, qui peuvent être produites par DIAL, PROSO et bien évidemment par SYNSEM.

Pour chacun des flux d'informations que nous venons d'évoquer, nous définissons ici ce que nous appelons une HYPOTHESE. Il s'agit d'une portion d'information élémentaire

qui transite entre deux modules, correspondant, pour le module émetteur, à un traitement particulier (étude d'une théorie, frontière de mot ou de syntagme à vérifier, nouvel énoncé dans un dialogue etc...).

Alors que nous détaillerons l'origine et la forme des hypothèses lexicales dans le prochain paragraphe, nous pouvons ici voir succinctement la forme que peuvent prendre les hypothèses syntaxico-sémantiques. Les hypothèses en provenance de PROSO peuvent être (dans l'absolu) de deux types : frontières de syntagmes ou modèles de structure syntaxique (affirmation, interrogation...). Nous pouvons donc ramener ces hypothèses à des objets de la forme :

{structure syntaxique} X [Intervalle de temps]

où l'intervalle de temps est la portion de l'énoncé où la structure hypothétisée est susceptible d'apparaître.

DIAL, quand à lui, produit des hypothèses qui peuvent être soit des structures syntaxiques, dans le cas où, après une question de la machine, l'utilisateur va probablement répondre par une phrase affirmative, soit une structure sémantique que DIAL peut avoir déterminée en fonction du contexte local d'analyse. Dans tous les cas, l'hypothèse doit être accompagnée d'une plage temporelle où elle est effective, ce qui permet de généraliser la forme que peut prendre une hypothèse syntaxico-sémantique en général :

{structure syntaxico-sémantique} X [Intervalle de temps]

C'est sous cette même forme que SYNSEM va fournir ses résultats d'analyse à DIAL, ce qui permet en particulier à celui-ci d'être averti que la reconnaissance de l'énoncé courant n'a été que partielle, et donc de réagir en conséquence.

Dans le cadre du travail présenté nous n'avons pas traité particulièrement les hypothèses syntaxico-sémantiques. Complémentaires par rapport aux hypothèses lexicales, nous avions pensé dans un premier temps que la réflexion effectuée sur ces dernières serait facilement généralisable. La stratégie de recherche adoptée ici est en effet du type *diviser pour régner*, ce qui suppose la possibilité de réunir les différentes analyses en une seule, sous une forme commune. Nous verrons, après avoir étudié les problèmes lexicaux, pourquoi cette réunification n'est pas nécessairement possible.

1.3 Etude d'un exemple : gestion des hypotheses lexicales.

1.3.1 Objectifs de cette étude.

Comme les hypothèses syntaxico-sémantiques que nous venons de présenter, les hypothèses lexicales peuvent prendre différentes formes. Tout d'abord, PROSO fournit à LEX des frontières de mots sous la forme de marques temporelles. Ces frontières contiennent en particulier les frontières de syntagmes fournis à SYNSEM, puisqu'une entrée lexicale ne fait jamais simultanément partie de deux groupements syntaxiques. Bien plus, les frontières de syntagmes sont beaucoup mieux détectées car elles marquent en général le rythme de l'énonciation (surtout en anglais [Lea 80], mais aussi en français), alors que les simples frontières de mots sont souvent masquées par des phénomènes de coarticulation.

LEX fournit des hypothèses de type mot, qu'il a effectivement vérifiées sur le treillis phonétique et qui sont donc parfaitement situées au niveau temporel. Ces hypothèses vont de manière exclusive à SYNSEM, afin d'être intégrées dans la construction d'une théorie. L'envoi possible d'hypothèses directement de LEX à DIAL, sujet de débat au sein de l'équipe, ne me semble pas pertinent, puisque DIAL n'est en mesure de comprendre que des formes sémantiques (en particulier pour les prédicats). Ainsi, si la médiocrité du signal à analyser oblige le système à fonctionner par mots-clefs, il est nécessaire pour chaque mot de passer par SYNSEM, afin d'acquérir sa propre structure sémantique. Quoi qu'il en soit, la production par LEX de telles hypothèses provient en fait de deux processus différents:

- Détermination ascendante de mots, directement sur le treillis phonétique. Les mots ainsi proposés possèdent alors une très bonne représentation acoustique et peuvent donc servir de points d'ancrage pour de nouvelles analyses syntaxico-sémantiques.

- Résultat de la vérification d'une hypothèse descendante; auquel cas il s'agit de mots correspondant à un état d'analyse bien précis, que LEX a pu reconnaître avec un score phonétique acceptable.

Enfin, SYNSEM et DIAL fournissent des prédictions lexicales descendantes, qui ne sont pas nécessairement de type mot, mais qui peuvent désigner des sous-lexiques[1] importants suivant le degré de finesse des analyses en cours au sein de ces deux modules. Ce sont ces hypothèses qui vont plus spécialement nous intéresser ici, puisqu'elles reposent sur des informations essentiellement structurelles et conceptuelles beaucoup plus que sur des renseignements tirés du signal.

A ce stade nous pouvons fournir un format standard pour toutes les hypothèses lexicales, qui combine les différents aspects présentés ci-dessus:

{Sous-lexique} X [plage temporelle]

où la plage temporelle correspond, soit au lieu où un sous-lexique prédit doit être vérifié dans le cadre d'une analyse descendante, soit à l'intervalle où un mot a été effectivement reconnu de façon ascendante (on aura remarqué qu'un mot est un sous-lexique particulier).

Dans ce cadre, où des hypothèses lexicales semblent transiter un peu dans tous les sens, quel va être notre objectif ? Nous ne pouvons rien faire des hypothèses ascendantes de LEX à SYNSEM, puisque celles-ci vont être directement intégrées dans des analyses nouvelles ou en cours. De même, les hypothèses de type frontières de mots vont surtout servir de confirmation ou tout simplement d'aide pour les mises en correspondance de mots avec le treillis phonétique. Une utilisation différente serait possible, mais elle impliquerait une gestion des problèmes temporels qui, à ce stade, n'est pas encore de mise.

Il nous reste donc les hypothèses en provenance de SYNSEM et de DIAL à traiter afin de générer un ensemble de mots pouvant être vérifiés au niveau phonétique. Comment réaliser cette fonction au niveau lexical ? Nous décrivons ici une proposition envisageable pour gérer ces traitements.

Tout d'abord, nous allons définir ce que peut attendre un module qui envoie une hypothèse à LEX. Deux cas se présentent, dont nous avons déjà plus ou moins parlé :

- le module émetteur est en train de travailler sur une certaine analyse, et de ce fait, il est arrivé à un point où il doit vérifier l'existence d'un (ou plusieurs) mot particulier au

[1] Par *sous-lexique*, on entendra tout ensemble de de mots extrait du lexique tout entier et utilisé dans un système parce qu'il possède une unité syntaxique, sémantique etc. ...

niveau du signal. Il s'agit donc d'*hypothèses de validation*. Typiquement, quand SYNSEM effectue une analyse syntaxique descendante et qu'il arrive à un nœud terminal de sa grammaire, il doit trouver un mot candidat qui convienne à cet endroit ; il génère alors une hypothèse correspondant au sous-lexique des mots qui vérifient les contraintes syntaxiques résultant de cette analyse.

- le module émetteur a réalisé une analyse satisfaisante à un niveau du dialogue (syntagme, phrase, énoncé, sous-dialogue) qui lui donne la certitude que certains mots vont apparaître sur les éléments du discours à venir. Il va donc en informer LEX afin que celui-ci soit guidé lors de ses recherches lexicales ultérieures. Ce sont alors des *hypothèses informatives*. Nous verrons au paragraphe suivant l'origine possible de telles hypothèses.

Pour pouvoir être comparées entre elles, les hypothèses informatives doivent être accompagnées d'un score qui représente l'importance que va prendre chacune d'entre elles quant aux choix pour une analyse ultérieure du système. Ces scores ne doivent pas être confondus avec d'éventuels scores de validation d'un mot sur le signal, mais plutôt être interprétés comme la confiance que met un module dans l'information transmise à LEX.

En fonction de toutes ces contraintes, nous pouvons commencer à définir un vrai gestionnaire d'hypothèses lexicales dont la position par rapport aux autres modules est schématisée dans la figure 1.2.

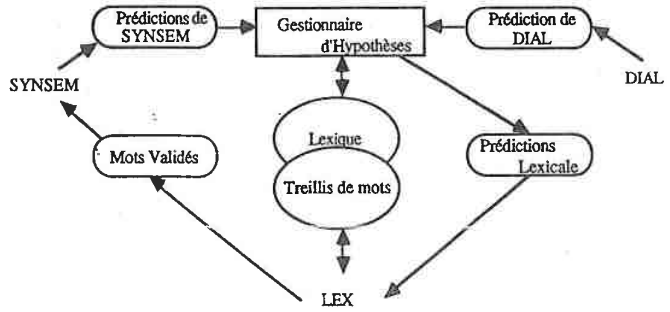


Figure 1.2.

Nous observons ici quelque chose d'important, à savoir que les informations manipulées par LEX sont de deux natures. Il y a d'un coté les informations lexicales proprement dites, c'est à dire l'état de connaissance de LEX à un instant donné sur l'ensemble des mots qu'il est susceptible de manipuler ; et d'un autre côté un treillis de mots, formé des lexèmes que LEX a déjà cherché à reconnaître à un endroit précis du treillis phonétique. Le rôle du gestionnaire d'hypothèses lexicales peut alors être défini comme une mise à jour du lexique dans le cas d'hypothèses informatives, ou une sélection de mots candidats pour la vérification, en vue de construire le treillis lexical.

1.3.2 Origine des hypothèses dans un dialogue.

Avant de détailler la mise en œuvre effective du gestionnaire d'hypothèses, nous allons faire un rapide aperçu de l'origine possible des hypothèses informatives qui arrivent à celui-ci. En effet, du dialogue jusqu'au niveau d'un énoncé, il est possible de trouver des indices dans les mémoires de travail de chacun des modules émetteurs qui permettent de générer des hypothèses pour LEX.

1 AVANCEMENT DU DIALOGUE.

A partir du modèle de gestion du dialogue utilisé par DIAL, de nombreuses prédictions peuvent être effectuées pouvant concerner le lexique. En début de dialogue, DIAL va ainsi prédire un énoncé de salutation en réponse à une introduction de la machine du type "Centre de renseignements administratifs, Bonjour!". Puis un sous-lexique particulier relatif à une demande d'informations va pouvoir être proposé (typiquement : "Je voudrais..."), correspondant à l'expression de la requête de l'usager. Enfin, après satisfaction de celui ci, on peut s'attendre à deux types d'énoncés auxquels correspondent deux sous-lexiques particuliers, à savoir : un énoncé de relance (une nouvelle question de l'usager) ou de fin de dialogue ("Je vous remercie, au revoir").

Ces connaissances générales au niveau d'un dialogue peuvent s'affiner à chaque échange puisque, progressivement, le thème général du discours se définit, comme par exemple une demande d'informations relative au renouvellement d'une carte d'identité, ou l'accession à la nationalité française. Localement, seul un sous-ensemble restreint du lexique relatif à ce thème peut être mis en exergue. On autorise néanmoins l'usage de mots moins probables, en cas de rupture de séquence, ou d'énoncé concernant la gestion du canal de communication, si l'usager désire préciser une réponse de la machine ("Pouvez vous répéter, s'il vous plaît"). Plus précisément, l'usage de certaines structures ou

expressions par le locuteur peut être fortement conditionné par un énoncé particulier de la machine ; comme dans le cas d'une question telle que :

- Où habitez vous ?

où la structure standard de réponse est du type :

- J'habite à [Nancy]

ou de façon plus elliptique :

- à [Nancy].

Cette grande corrélation entre deux énoncés successifs permet de restreindre de manière importante l'espace d'analyse au niveau du lexique, et bien sûr, de façon plus générale les structures de la langue.

2 INFORMATIONS FOURNIES PAR SYNSEM.

Les prédictions plus précises au niveau de la structure détaillée d'un énoncé ne concernent plus le module de dialogue, mais plutôt les deux modules d'analyse syntaxico-sémantique et de prosodie. Ce dernier processeur fournit à l'analyseur lexical des frontières possibles de mots, sans être en mesure de préciser la nature exacte de ces mots. Les hypothèses fournies sont donc purement temporelles. Les analyseurs peuvent, quant à eux, utiliser des résultats relativement sûrs, obtenus à partir de la reconnaissance d'une partie d'énoncé, pour induire des hypothèses sur les éléments restants. Si par exemple le système a reconnu la portion d'énoncé :

J'habite ...,

une hypothèse sémantique va pouvoir être générée concernant un sous-lexique de lieu. Ceci peut se faire dans la pratique grâce à des contraintes sémantiques sur les constructions possibles autour du verbe "habiter". Ce phénomène de ***déclenchement sémantique*** ('semantic priming') est d'ailleurs connu en psychologie dont les résultats apportent beaucoup à notre approche [Heyer 85].

1.3.3 Définition d'un gestionnaire d'hypothèses.

Les différentes informations dont nous venons de signaler l'existence vont donc pouvoir mettre en évidence certains sous-lexiques particuliers afin de faciliter la tâche de vérification lexicale impartie à LEX. Nous répétons que notre but n'est pas de couper le lexique en deux, avec d'un côté les mots susceptibles d'apparaître, et de l'autre, tous les autres, qui se retrouvent mis à l'écart. Dans un dialogue ***naturel***, il est toujours acceptable que des changements brusques de thème apparaissent en fonction des lubies de l'usager,

ou simplement parce qu'il a mal compris la dernière intervention du système ; auquel cas on entre dans un nouveau sous-dialogue. Si les prédictions faites ne conviennent pas à la situation de dialogue courante, et que le système se retrouve de ce fait incapable d'analyser l'énoncé à venir (ou pire s'il reconnaît ce qu'il a prédit plutôt que ce qui existe vraiment), nous n'aurons pas atteint notre objectif. Nous devons donc concevoir un gestionnaire d'hypothèses lexicales qui traite bien toutes les hypothèses en provenance des autres modules, et qui, par un calcul approprié, gère les différents scores correspondants, afin d'en répercuter l'importance toute relative au niveau de la base de données lexicales dont il dispose.

Nous voyons, d'après ces quelques remarques, se dessiner la méthode que nous avons adoptée pour mener à bien les objectifs exprimés plus haut. La définition d'un gestionnaire d'hypothèses lexicales peut s'appuyer sur deux aspects à la fois distincts et intimement liés :

- la proposition d'une structure pour les informations lexicales qui permette d'exprimer de façon simple et efficace les sous-lexiques présents dans les hypothèses;
- la mise en œuvre d'un calcul de score possédant les bonnes propriétés sus-citées, mais surtout qui s'adapte parfaitement au choix de structure effectué.

Nous allons maintenant étudier ces deux aspects successivement, en justifiant les différents choix qui ont été faits par comparaison avec d'autres alternatives, mais surtout en montrant qu'il est possible de construire ainsi une unité de gestion cohérente et autonome.

1.3.4 Choix d'un calcul de score.

1 QUEL CALCUL DE SCORE ?

Le choix d'un calcul de score est une opération particulièrement difficile quand on regarde le nombre de travaux effectués sur ce sujet. De nombreuses mesures de l'incertain sont effectivement utilisables pour combiner des informations approximatives provenant de sources de connaissances variées. Pour un inventaire structuré de la plupart de ces mesures, on peut se référer aux travaux de Dubois et Prade ([Prade 82], [Dubois 80] ou [Dubois 86]). Sans entrer dans les détails, nous pouvons appuyer notre choix en comparant deux grands classiques du calcul de l'incertain : la théorie des possibilités et le calcul des probabilités.

La théorie des possibilités repose initialement sur les travaux de Zadeh (65) concernant les ensembles flous, et a rapidement connu une popularité importante, tant dans les milieux de l'Intelligence Artificielle que de la Linguistique. En effet, le concept d'ensemble flou a pour but d'étendre la notion mathématique de fonction caractéristique d'un ensemble, afin de pouvoir exprimer certains prédicats "flous" utilisés dans la vie courante. On définit donc un sous-ensemble de l'espace des objets manipulés Θ, par une fonction μ, de Θ dans [0,1], sensée refléter le degré d'appartenance de l'objet au sous-ensemble flou considéré. On peut par exemple représenter le concept associé à l'adjectif "grand" (pour un être humain, masculin, occidental...) sur l'axe réel de la façon suivante:

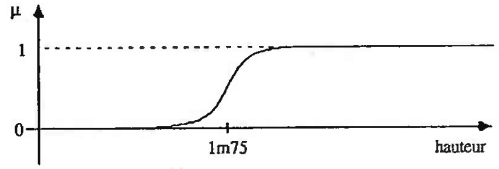


Figure 1.3.

A partir de cette fonction d'appartenance on peut définir deux mesures de possibilité et de nécessité Π et N qui permettent d'évaluer le degré d'inclusion ou d'intersection d'un ensemble classique dans l'ensemble flou représenté par μ :

- $\Pi(X) = \max_{x \in X} \mu(x)$.
- $N(X) = \min_{x \in X} \mu(x)$.

Une autre façon de définir ces deux mesures repose sur une suite de sous-ensembles de l'espace de référence (les réels dans l'exemple ci-dessus), croissante au sens de l'inclusion, appelés α-coupes ([Dubois 86]). Si $(A_\alpha)_{0 \le \alpha \le 1}$ est une telle suite, on observe les propriétés suivantes :

- $A_\alpha = \{ \omega \in \Theta \text{ tel que } \mu(\omega) \ge \alpha \}$
- $\forall \omega \in \Theta, \mu(\omega) = \sup \{ \alpha \text{ tel que } \omega \in A_\alpha \}$.

Nous verrons par la suite l'intérêt d'une telle définition en étudiant la théorie de Dempster-Shafer.

Le principal reproche qui peut être fait à la théorie des possibilités vient de l'usage sans partage qui est fait des opérations mathématiques 'min' et 'max'. En effet, chaque agrégation d'information sur un corpus déjà existant fait systématiquement perdre une certaine quantité de connaissances, ce qui interdit en particulier tout retour en arrière en cas d'erreur de la part de la source ayant apporté ces informations.

A l'opposé, le calcul des probabilités ne peut pas encourir de tels reproches. Il se caractérise en effet par une grande précision dans ses évaluations de l'incertain, comme le rappelle P. Cheeseman (85) (on trouvera une excellente synthèse sur les probabilités dans Papoulis (65)). Pour un ensemble de distributions sur un espace de référence, on peut ainsi déterminer une nouvelle distribution probabiliste optimale dans tous les cas. Cependant, tous ces calculs reposent sur des probabilités conditionnelles dont la détermination demande toujours des calculs nombreux et complexes, souvent rédhibitoires pour nos moyens actuels. De plus, la nécessité d'une telle précision au sein d'un système de dialogue oral homme-machine ne se fait pas obligatoirement sentir, puisque nous cherchons avant tout à guider sommairement la reconnaissance d'un énoncé. Un autre argument que l'on peut avancer contre une telle précision, est que, de toute façon, nous ne disposons pas de données précises en entrée, que ce soit au niveau des scores de phonèmes qui sont calculés de manière assez intuitive, ou même au niveau des données extraites de corpus qui peuvent être extrêmement sensibles aux circonstances de l'expérimentation.

Entre ces deux extrêmes de flou et de précision, nous avons choisi d'utiliser une théorie parallèle, celle de Dempster-Shafer. En fait, elle peut surtout être située à un niveau plus élevé de généralité par rapport aux possibilités et aux probabilités. Il est ainsi possible de retrouver certains résultats de ces deux théories, en fonction de la structure adoptée pour les distributions, au sens de la théorie de Dempster.

2 GÉNÉRALITÉS SUR LA THÉORIE DE DEMPSTER-SHAFER.

Contrairement à certaines approches axiomatiques considérées par les auteurs que nous venons de citer, cernant une mesure d'incertitude par les propriétés qu'elle doit satisfaire, Dempster, puis son élève Shafer (76) définissent une répartition de notre connaissance parmi certains sous-ensembles particuliers de l'espace de référence.

Nous appelerons Θ un espace d'objets quelconques, les mesures de confiance sur cet espace de référence s'expriment à partir d'une *pondération probabiliste de base* ('basic probability mass') : m ([Barnett 83]) vérifiant:

- $m(\emptyset) = 0$; aucune masse n'est allouée à l'ensemble vide.
- $\sum_{\theta \supset A} m(A) = 1$; la somme des masses allouées aux différents sous-ensemble de Θ est égale à l'unité (condition de normation).

Cette fonction m de 2^{Θ} (ensemble des parties de Θ) dans [0,1] exprime la quantité de confiance que l'on alloue à un ensemble particulier. Chaque sous-ensemble de Θ pour lequel m prend une valeur non nulle est appelé *élément focal*. Ces éléments focaux peuvent se répartir à volonté dans Θ, et en particulier, on peut avoir $m(\Theta) \neq 0$. Cette dernière valeur exprime l'incertitude relative à l'ensemble tout entier, quand par exemple, on n'est pas sûr d'avoir recensé tous les éléments de l'ensemble de référence.

La fonction m peut être vue comme une masse que l'on affecterait à certains sous-ensembles de Θ et pour laquelle on ne connaît pas de répartition plus fine au sein de chacun d'eux. Lors de la définition des fonctions qui suivent, on pourra imaginer cette masse sous forme de 'gravillons' susceptibles de se déplacer à l'intérieur du sous-ensemble correspondant.

On définit les deux fonctions de Crédibilité et de Plausibilité de la façon suivante:

- Cr(X) exprime la confiance que l'on peut sûrement avoir vis-à-vis du sous-ensemble, c'est-à-dire :

$Cr(X) = \sum_{X \supset A} m(A)$, représente la totalité des masses contenues dans X.

- Pl(X) est la confiance maximale susceptible d'être affectée à X :

$Pl(X) = \sum_{A \cap X \neq \emptyset} m(A)$; On voit que Pl(X) est l'ensemble des masses qui peuvent être déplacées à l'intérieur de X, tout en restant dans leurs éléments focaux assignés.

Les deux fonctions Cr et Pl définies à l'instant forment, pour certaines pondérations m données, un couple de mesures de confiance *duales*, au sens établi par Prade (82). En particulier on retrouve dans certains cas particuliers des mesures que nous avions citées. Par exemple, si les éléments focaux forment une suite d'éléments emboîtés (monotone au sens de l'inclusion), les mesures de Crédibilité et de Plausibilité sont respectivement des mesures de Nécessité et de Possibilité associées à min et max.

Au contraire, si les éléments focaux forment une partition de l'espace Θ, les mesures de Crédibilité et de Plausibilité deviennent probabilistes et expriment, par leur différence, l'incertitude quant à la répartition des masses dans chaque élément focal. En particulier, si on désire définir une mesure de probabilité 'p' compatible avec la pondération m, c'est-à-dire telle que:

$\forall \Theta \supset A$, $(m(A) \neq 0) \Rightarrow (p(A) = m(A))$, celle-ci se trouve obligatoirement comprise entre Cr et Pl.

La plupart des auteurs décrivent l'encadrement de mesures de probabilités comme étant idéal pour exprimer l'incertitude relative à la probabilité effective d'un événement. Ces considérations ne sont pourtant pas toujours constructives et s'éloignent de la réalité des probabilités (au sens d'une estimation). La 'bonne' solution (si j'ose employer cet adjectif!) semble être d'interpréter la pondération de base en terme d'entropie maximale [Oswald 86]. On exprime donc la pondération probabiliste la plus générale possible (celle qui n'apporte aucune information, hormis celle qui est disponible) en répartissant uniformément les masses sur chaque singleton de l'élément focal correspondant. Cette approche est particulièrement encouragée dans [Cheeseman 85].

Enfin le cas le plus spécifique du précédent apparaît quand les deux mesures de Crédibilité et de Plausibilité sont égales. Les éléments focaux sont alors réduits aux singletons de l'espace Θ, et la pondération 'm' est une véritable *distribution de probabilité.*

Un problème cependant demeure, que l'approche de Shafer ne résout pas. En effet, aucune indication ne permet, à moins d'étudier avec précision la pondération probabiliste de base, de trouver un mode de combinaison des fonctions Cr et Pl de manière interne. Il est nécessaire de revenir systématiquement à la pondération pour recalculer les deux mesures d'incertain à partir des définitions.

Malgré tout, cette théorie est particulièrement utile dans le domaine qui nous concerne, puisqu'elle offre un moyen de combiner diverses sources d'information, grâce à la règle de Dempster dont nous définissons ici les principaux aspects.

On considère que l'on dispose de deux sources d'information, qui fournissent leurs données sous la forme de deux pondérations probabilistes de base, m1 et m2. Ces pondérations font ressortir deux suites d'éléments focaux $(A_i)_{1\le i\le n}$, et $(B_j)_{1\le j\le m}$ pour lesquelles on suppose qu'il existe un i et un j particuliers tels que $A_i \cap B_j \neq \emptyset$.

On définit l'opération de combinaison $m = m1 \oplus m2$, qui fournit une nouvelle pondération s'exprimant grâce aux formules suivantes:

- $\forall \Theta \supset A,\ m(A) = K \cdot \sum_{X \cap Y \neq \emptyset} m(X).m(Y)$

- avec $K = \dfrac{1}{1 - \sum_{X \cap Y = \emptyset} m(X).m(Y)}$

K est un facteur de normation, qui permet à la nouvelle pondération d'avoir une masse totale égale à 1. La formule ci-dessus explique la restriction imposée aux deux masses m1 et m2 : posséder au moins deux éléments focaux non-disjoints. En effet, l'opération de combinaison de Dempster a pour effet de renforcer la masse allouée aux sous-ensembles que corroborent simultanément m1 et m2, alors que les sous-ensembles qui ne sont désignés que par une seule des deux pondérations se voient affecter une nouvelle masse nulle. Ainsi, si m1 et m2 ne soutiennent simultanément aucun élément de Θ, le facteur K devient infini, et l'opération n'est plus définie.

Pour bien comprendre le fonctionnement de l'opération de combinaison, on peut la représenter sur le schéma qui suit. On considère ici deux pondérations m1 et m2 d'éléments focaux respectifs (A,B) et (X,Y,Z). Le long de deux des cotés adjacents du carré on a représenté chacune des pondérations en juxtaposant les masses élémentaires allouées à chaque élément focal. La longueur totale est donc de 1 , et la surface du carré est alors elle-même égale à l'unité, ce qui représente toute la masse qu'il est possible de redistribuer par combinaison. Chaque zone particulière, que le découpage fait ressortir est une masse (sa surface) qui sera affectée à l'intersection des deux éléments focaux correspondant.

Par exemple, si $B \cap Y = \emptyset$, m(B).m(Y) servira au calcul de K, sinon, $B \cap Y$ sera un élément focal de la pondération m, et une partie de sa masse sera m(B).m(Y).

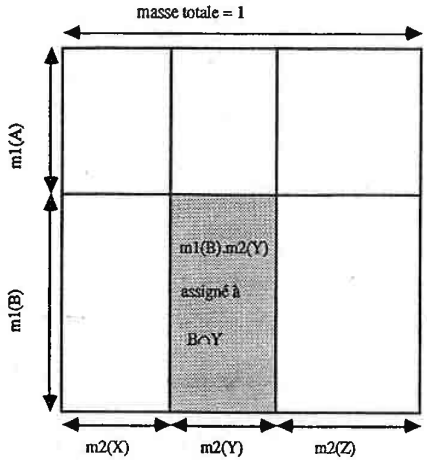


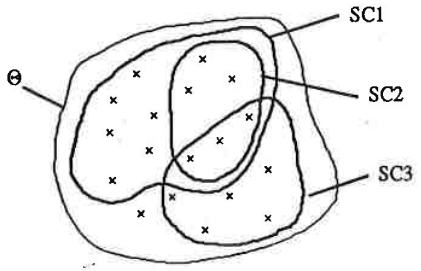
Figure 1.4 : combinaison de deux distributions d'après Shafer (76).

Reposant sur les opérations d'intersection ensembliste et de somme ou de produit sur les nombres, la combinaison de Dempster possède de nombreuses propriétés intéressantes. Elle est commutative et associative, donc l'ordre de combinaison de plusieurs sources de connaissance n'influe pas sur le résultat. De deux pondérations probabilistes de base, on obtient une nouvelle pondération et il est alors possible de redéfinir des mesures de Crédibilité et de Plausibilité. Cependant, pour une structure initiale des pondérations, on n'est pas sûr d'obtenir cette même structure au niveau du résultat de la combinaison. En particulier, si on possède deux pondérations possibilistes définies à partir d'α-coupes (éléments emboîtés), il faut généralement que la réunion des deux suites d'éléments focaux puisse être ordonnée au sens de l'inclusion, pour retrouver une pondération possibiliste après combinaison (les deux sources désignent le même résultat au sens d'un calcul d'erreur).

Il est étonnant de constater que le cas probabiliste est particulièrement compatible avec la règle de combinaison de Dempster. En effet, si les éléments focaux de m1 et m2 forment deux partitions de l'espace, la pondération résultat se répartit à son tour sur une nouvelle partition, éventuellement plus fragmentée, de Θ. L'apport d'information a donc permis d'*affiner le résultat.*

3 APPLICATION AU LEXIQUE.

Dans le cadre défini ci-dessus, le lexique peut être vu comme l'espace de référence Θ, dont les éléments sont tout simplement les entrées lexicales. En effet, il est possible d'assimiler les hypothèses informatives qui arrivent au gestionnaire à un ensemble de sources de connaissances mettant chacune en avant un certain sous-lexique, comme cela est schématisé dans la figure 1.5.



SCi : Sources de connaissance.

Figure 1.5 : apport de différentes sources de connaissances sur un espace de référence.

Suivant le formalisme de Dempster-Shafer, nous pouvons donc assimiler une hypothèse à une distribution de masse sur le lexique Θ, dont un élément focal est le sous-lexique désigné, que nous nommerons par la suite X. Il nous reste donc à déterminer les autres éléments focaux de la distribution. Or, on sait que le module émetteur d'une hypothèse désire informer LEX que les mots du sous-lexique X vont *plus probablement* apparaître sur le signal, ceci avec une incertitude donnée, et sans fournir aucune information sur le reste du lexique. Par conséquent, afin de conserver une cohérence aux informations échangées entre les modules, si on suppose avoir alloué une masse q à X, la masse restante (1-q) ne peut être qu'allouée à ¬X ou à Θ et correspond alors à l'évaluation de l'incertitude que le module émetteur confère à son hypothèse. Nous obtenons donc deux modélisations possibles pour une hypothèse lexicale sous la forme d'une distribution de masse :

- Deux éléments focaux X et ¬X affectés respectivement des masses q et 1-q, que nous noterons (X; ¬X; q).
- Deux éléments focaux X et Θ affectés respectivement des masses q et 1-q, que nous noterons (X; Θ; q).

Nous avons montré ([Romary 87]) que ces deux distributions étaient équivalentes, puisque pour toute hypothèse (X; ¬X; q), il existe une hypothèse (X; Θ; q'), avec q'=

$\frac{1}{2-q}$, dont l'effet est le même par le calcul que nous allons développer ci-après. Nous ne conserverons dorénavant que la dernière forme, qui semble plus facile à appréhender intuitivement, puisque la masse non spécifiquement impartie à X est affectée au lexique dans son ensemble, sans plus de précision.

Nous devons maintenant modéliser le lexique lui-même comme une distribution de masse définie à partir d'éléments focaux particuliers. Dans un premier temps, nous pouvons observer que les seules données dont nous disposons pour initialiser une telle distribution sont réparties sur chaque entité lexicale. L'étude d'un corpus touchant le domaine du dialogue peut en effet nous donner un ensemble de fréquences d'apparition pour chacun des mots. Au sens de Dempster-Shafer, nous aurions donc une distribution probabiliste vraie, où les éléments focaux sont chacun des singletons du lexique. Il est tout de suite visible qu'une distribution de la sorte compliquerait beaucoup trop les calculs pour être admissibles.

Nous pouvons contourner ce problème grâce à la forme de la distribution adoptée pour l'hypothèse et une propriété qui en découle. En effet, il est possible de montrer que la combinaison de $m_h = (X; \Theta; q)$ (de l'hypothèse) avec la distribution envisagée équivaut à combiner m_h avec une distribution d'éléments focaux X et $\neg X$ affectés respectivement des masses p et 1-p, où p est la somme des masses contenues dans X initialement (à partir de fréquences par exemple). La nouvelle distribution, au niveau de chaque entité lexicale, est alors obtenue en répartissant le résultat de l'opération de combinaison proportionnellement à la distribution d'origine. La démonstration de cette propriété est donnée pour mémoire à la fin de cette section.

La distribution ainsi définie pour le lexique, notée $m_i = (X; \neg X; p)$, est facilement calculable, puisqu'il ne s'agit que de sommer un ensemble de fréquences sur un sous-lexique particulier. Nous pouvons maintenant expliciter l'opération d'agrégation d'une hypothèse au niveau lexical par la règle de Dempster-Shafer, qui donne une nouvelle distribution finale : m_f telle que ([Shafer 76], [Barnett 83]):

$$m_f(X) = K.p \; ; \; m_f(\neg X) = K.(1 - p).(1 - q)$$

Où $K = \frac{1}{1 - q + p.q}$ est un coefficient de normation, de sorte que :

$m_f(X) + m_f(\neg X) = 1$.

Cette combinaison peut alors être schématisée par le diagramme de la figure 1.6.

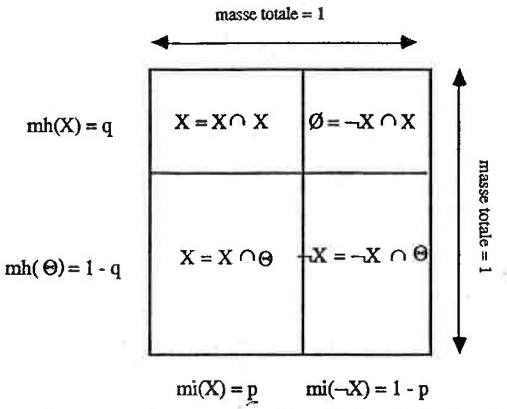


Figure 1.6 : combinaison d'une hypothèse à la distribution initiale.

Les propriétés de cette combinaison sont celles propres à la règle de Dempster, à savoir la commutativité et l'associativité, auxquelles peuvent être ajoutées des propriétés particulières dues aux distributions que nous avons utilisées.

Tout d'abord, nous pouvons obtenir une idée de la façon dont la combinaison agit sur un sous-ensemble donné du lexique, en observant la quantité représentant la différence entre les masses finales et initiales allouées à X :

$$\Delta = m_f(X) - m_i(X) = \frac{p.q.(1 - p)}{1 - q + p.q};$$

Δ est toujours positive pour des valeurs non extrêmes de p et de q, ce qui indique un réel renforcement de X, qui est plus important pour des valeurs élevées de q. La combinaison successive d'hypothèses est donc d'autant plus efficace que plusieurs corroborent. Ainsi, si les modules DIAL et SYNSEM détectent des indices favorables à un sous-lexique relatif à l'expression d'un lieu, celui-ci sera beaucoup plus renforcé.

Comme cela avait été indiqué par Barnett (83), Ginsberg (84), ainsi que par Lu et Stephanou (84), il peut être possible de recouvrer la distribution d'origine à partir de celle calculée, dès lors que les éléments focaux utilisés ont de bonnes propriétés. Dans le cas présent, nous pouvons montrer que toute combinaison d'hypothèses est réversible, et ceci indépendamment d'autres hypothèses qui auraient été reçues par le gestionnaire entre temps.

Ainsi, si l'hypothèse représentée par la distribution : { $m_h(X) = q$; $m_h(\Theta) = 1 - q$ } a été combinée avec la distribution d'origine : { $m_i(X) = p$; $m_i(\neg X) = 1 - p$ } pour donner la distribution m_f, alors, l'hypothèse correspondant à la distribution : { $m_h'(\neg X) = q$; $m_h'(\Theta) = 1 - q$ } combinée avec celle calculée : m_f, produit la distribution d'origine m_i. C'est un résultat important, pour un fonctionnement correct d'ensemble de l'architecture proposée, puisqu'un module ayant envoyé des informations vers le lexique peut s'apercevoir qu'il a commis une erreur, et se rétracter, pendant que les autres modules continuent d'envoyer des éléments de connaissance, sans qu'apparaissent pour autant des interférences.

Enfin, comme nous l'avions signalé, il est possible, après application de la combinaison d'une hypothèse, de recalculer facilement une distribution de masse au niveau de chacun des singletons de Θ. Si nous appelons Mf cette distribution fine, la nouvelle masse $M(f\{\omega\})$ allouée au singleton $\{\omega\}$ est immédiatement fournie par multiplication de l'ancienne valeur $Mi(\{\omega\})$ par un facteur particulier suivant que ω appartient à X ou à $\neg X$. D'où, si la règle de Dempster a donné comme résultat :

$m_f(X) = K . p$; $m_f(\neg X) = K . (1 - q) . (1 - p)$,

nous obtenons comme nouvelle distribution 'précise', Mf, associée à la distribution grossière, 'm_f' :

$Mf(\{\omega\}) = K . Mi(\{\omega\})$, si $\omega \in X$, et

$Mf(\{\omega\}) = K . (1 - q) . Mi(\{\omega\})$, si $\omega \in \neg X$.

La nouvelle distribution est alors probabiliste, elle reste consistante et peut reconduire à la distribution d'origine par le mécanisme décrit plus haut.

DÉMONSTRATION DE L'ÉQUIVALENCE DE NOTRE CALCUL AVEC LE CALCUL DIRECT

Soit Mi la distribution d'origine (probabiliste) répartie sur tous les singletons $\{\omega\}$, qui en forment alors les éléments focaux. Soit $m_i = (X; \neg X; p)$, la distribution obtenue en sommant sur X et $\neg X$ les masses totales qu'ils contiennent respectivement. Nous avons donc :

$m_i(X) = p$

$m_i(\neg X) = 1 - p$; puisque la distribution est normée.

Soit $m_h = (X; \Theta; q)$ la distribution correspondante à l'hypothèse lexicale à combiner avec Mi, on a :

$m_h(X) = q$

$m_h(\neg X) = 1 - q$

→ Montrons alors que le résultat obtenu : Mf est le même si on combine directement m_h et Mi ou si on combine m_h et m_i, avant de répartir le résultat sur X et $\neg X$ proportionnellement à la distribution d'origine Mi.

Soit $(A_i)_{1 \le i \le n}$ les singletons tels que $X \supset A_i$, et $(B_j)_{1 \le j \le m}$ les singletons tels que $X \supset B_j$. Le calcul direct (schématisé par la figure 1.7 pour Mf en combinant m_h et Mi donne:

- $\forall$ i, $m_f(A_i) = K . (Mi(A_i).m_h(X) + Mi(A_i).m_h(\Theta))$; ce qui donne sachant que $m_h(X) + m_h(\Theta) = 1$:

$$m_f(A_i) = K . Mi(A_i)$$

- $\forall$ j, $m_f(B_j) = K . Mi(B_j).m_h(\Theta) = K . Mi(B_j) . (1 - q)$

- avec $K = \dfrac{1}{\sum_{i=1}^{n} Mi(A_i) + \sum_{i=1}^{m} Mi(B_j).(1-q)}$; c'est à dire :

$$K = \frac{1}{p + (1-p).(1-q)} = \frac{1}{1 - q + pq}$$

Ce qui est bien équivalent aux résultats que nous avions obtenus précédemment en passant par l'intermédiaire de la distribution m_i. Notre méthode, plus efficace par ailleurs sur le plan algorithmique donne donc un résultat correct.

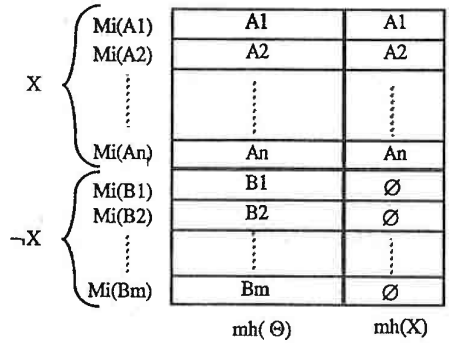


Figure 1.7 : Combinaison des distributions Mi et m_h.

1.3.5 Définition d'une structure adaptée.

Une fois défini un calcul de score permettant, formellement, de prendre en compte l'influence d'une hypothèse particulière au niveau du lexique, il nous reste à définir une structure pour celui-ci qui permette de façon effective la mise en œuvre des besoins nécessaires pour ce calcul. Ces besoins sont de deux ordres :

- le lexique doit être accessible pour les autres modules à partir des informations dont ils disposent. La réalisation de ce besoin peut passer par une structuration du lexique à partir des traits syntaxiques ou sémantiques utilisés par SYNSEM et DIAL. Ainsi, ces derniers pourraient sélectionner un sous-lexique en fournissant une fonction booléenne de ces traits qui le désignerait. Par exemple, si on utilise les traits manipulés par notre grammaire de cas, la formule *inter(nom +mvt)* va désigner tous les noms prédicatifs indiquant un mouvement comme "départ", "retour", "déplacement" etc. Nous utilisons ainsi les opérations 'inter', 'union' et 'non' (complémentaire d'un ensemble).

- les calculs présentés pour combiner une hypothèse à une distribution de masse sur le lexique doivent pouvoir être faits dans les meilleures conditions, c'est à dire suivant des algorithmes optimalisés par rapport à la structure du lexique.

Un dernier point à signaler concerne le statut du lexique dans un système oral de compréhension de la parole. Jean Marie Pierrel a déjà signalé dans sa thèse (81) que le lexique était un point de convergence de nombreuses informations et qu'il devait donc être construit de telle sorte qu'il soit accessible effectivement en fonction de chacune d'entre elles.

En fonction de ces différentes contraintes, nous avons choisi une structure du lexique sous la forme d'une arborescence multiple sur la base des traits qui s'imposent, étant donnés les modèles linguistiques utilisés au niveau de SYNSEM et de DIAL. Nous considérons donc que les mots du lexique forment les feuilles d'un certain nombre de graphes acycliques orientés. A une certaine catégorie de traits, on associe ainsi un graphe qui permet d'accéder au lexique sous un angle particulier.

Nous allons expliciter ces notions d'arborescence à partir de l'étude de l'exemple correspondant à la figure 1.8.

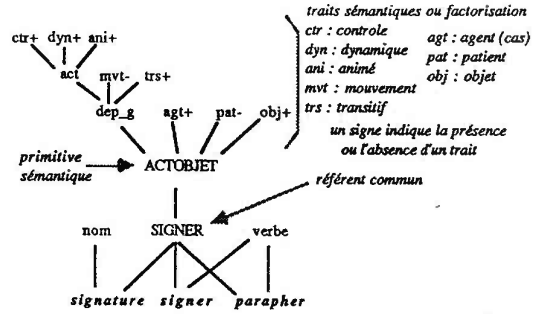


Figure 1.8

Dans cet exemple nous pouvons voir trois mots du lexique, 'signature', 'signer' et 'parapher' qui sont associés tous trois à une même forme sémantique : 'SIGNER' (le concept général de signer). 'SIGNER' entre lui-même dans le cadre d'une primitive sémantique utilisée par la grammaire de cas : 'ACTOBJET'. Enfin, cette primitive est déterminée par un certain nombre de traits formant ses ancêtres, soit directement, soit par factorisation en un nœud intermédiaire.

La deuxième arborescence présentée figure 1.8 repose sur des traits syntaxiques qui peuvent être plus complexes (type de conjugaison pour les verbes, genre pour les noms, etc). Elle montre en particulier que les différentes structurations sont disjointes pour tous les nœuds hormis les feuilles de chaque arbre. Les différents accès au lexique sont donc effectivement séparés, mais peuvent être combinés pour accéder à un sous-lexique.

La structure précédente est intéressante parce qu'elle permet d'effectuer efficacement les opérations nécessaires au gestionnaire d'hypothèses que nous avons défini. En effet, l'accès à un sous-lexique est réalisé par l'intermédiaire des différents nœuds des arborescences à partir desquels on peut désigner les mots du lexique par des techniques de propagation de marqueurs. Ces techniques, qui s'adaptent particulièrement bien à une programmation parallèle, sont généralement utilisées pour la gestion des mécanismes d'héritage ([Touretzky 86]).

Les marqueurs sont un ensemble d'étiquettes affectées aux nœuds des différentes arborescences. Chaque nœud peut posséder plusieurs marqueurs. Dans notre application, un marqueur va représenter un sous-lexique particulier désigné par tout ou partie d'une formule booléenne sur les nœuds (ou les traits associés). A chaque opération booléenne

va donc être associée une opération sur les marqueurs qui récursivement, permettra de calculer des formules complètes conformément aux algorithmes présentés ci-dessous.

Soit $\mathcal{N}$ un ensemble de nœuds d'une arborescence et $\mathcal{M}$ un ensemble infini de marqueurs (autant que nécessaire pour les calculs). On définit tout d'abord quelques fonctions de travail.

- **nouveau_marqueur()** donne un nouvel élément de $\mathcal{M}$, c'est à dire un marqueur non encore utilisé.

- pour n_0 un nœud de $\mathcal{N}$ et μ un marqueur de $\mathcal{M}$, **indique_marqueur**(n_0,μ) attache le marqueur μ au nœud n_0 et **rappel_marqueur**(n_0,μ) indique (par un booléen 'vrai' ou 'faux') si le nœud n_0 possède le marqueur μ.

- pour un nœud n_0 de $\mathcal{N}$, **fils_de()** est une fonction qui rend l'ensemble des nœuds de $\mathcal{N}$ qui sont des fils directs de n_0.

- pour un nœud n_0 de $\mathcal{N}$ et une fonctionnelle Λ (lambda) de domaine $\mathcal{N}$, **propage_fonction**(n_0,Λ) applique Λ_{n_0} à n_0 et s'appelle récursivement sur chacun des fils de celui-ci suivant l'algorithme suivant :

```
propage_fonction(n0,Λ)
début ;
N = fils_de(n0) ;
apply(Λ, n0) ;                    /* on applique la fonctionnelle à n0 */
pour n dans N
    propage_fonction(n,Λ) ;
fin ;  /* la boucle s'arrête quand il n'y a plus de nœud à explorer */
```

Remarque : les fonctionnelles seront indiquées sous forme de lambda expressions.

- pour une fonctionnelle Λ de domaine $\mathcal{N}$, **calc_general**(Λ) applique Λ à tous les nœuds de $\mathcal{N}$.

A partir de ces éléments, on peut donner la fonction **calc_formule**(F) calculant une formule booléenne F sur les nœuds d'une arborescence. Cette fonction renvoie le marqueur associé au sous-lexique désigné par F.

```
calc_formule(F)
début
si F appartient à N                /* on arrive sur un nœud directement */
    début ;
    μ= nouveau_marqueur();
    propage_fonction(F, λf.indique_marqueur(n,μ) ) ;
                                   /* on marque toute la descendance de F */
    fin ;
sinon si F = (non F1)
    début ;
    μ1 = calc_formule(F1) ;
    μ2 = nouveau_marqueur();
    calc_general(λn.{si  rappel_marqueur(n,μ1) alors
                         indique_marqueur(n,μ2) }
                 );
    return μ2 ;
    fin ;
sinon si F = (union F1 F2)
    début ;
    μ1 = calc_formule(F1) ;
    μ2 = calc_formule(F2) ;
    μ3 = nouveau_marqueur();
    calc_general( λn.{si   rappel_marqueur(n,μ1)
                               ou  rappel_marqueur(n,μ2)
                         alors
                               indique_marqueur(n,μ3) }
                 );
    return μ3 ;
    fin ;
sinon si F = (inter F1 F2)
    debut ;
    μ1 = calc_formule(F1) ;
    μ2 = calc_formule(F2) ;
    μ3 = nouveau_marqueur();
    calc_general( λn.{si   rappel_marqueur(n,μ1)
                               et  rappel_marqueur(n,μ2)
                         alors
                               indique_marqueur(n,μ3) }
                 );
    return μ3 ;
    fin ;
fin ;
```

Maintenant que nous savons marquer toute formule booléenne sur un ensemble de nœuds, l'application du calcul de score est immédiate, puisqu'il se résume (1.3.4) à une opération de combinaison au niveau d'un sous-lexique tout entier, suivie d'une multiplication du score de chaque mot par un facteur dépendant de son appartenance au sous-lexique désigné ou à son complémentaire. Or cette information est directement disponible à travers le marqueur qui a été calculé.

1.3.6 Implémentations et résultats.

Une maquette de ce gestionnaire a été réalisée en Flavor ([Moon 86]), un langage orienté objet dont nous avons développé notre propre version en Lelisp 15.1 ([Chailloux 86]). Ce type de langage s'adapte parfaitement bien à la représentation d'une structure hiérarchique telle que celle que nous avons envisagée pour notre lexique. Cependant,

comme les contraintes exigées d'une relation d'héritage au niveau d'une application particulière sont toujours différentes, de nombreuses variantes par rapport au langage support sont à introduire et dans les cas les plus extrêmes, la relation doit être redéveloppée entièrement. Néanmoins, un langage orienté objet est toujours idéal pour développer une telle maquette de façon très structurée et très facile à maintenir.

Le fait d'avoir réécrit entièrement un noyau de langage a finalement permis de l'adapter spécifiquement à notre problème. En particulier, nous avons développé des fonctionnalités qui simulent une implémentation parallèle des algorithmes de calcul de score et d'opérations ensemblistes. Ces techniques reposent sur deux fonctions que nous avons déjà rencontrées : **calc_formule**() et **propage_fonction**(), qui toutes deux permettent d'appliquer une fonctionnelle à un ensemble de nœuds.

Dans notre implémentation, chaque nœud de l'arborescence est un *flavor* du langage (une classe) dont une instance particulière assure la gestion. Les seuls flavors qui peuvent posséder plus d'une instance sont bien évidemment ceux qui sont associés aux mots du lexique et dans ce cas, ces instances représentent les différentes occurrences du mot donné sur le treillis phonétique. La structure de donnée utilisée est donc parfaitement isomorphe à la structure formelle que nous avions envisagée pour le lexique, ce qui permet par exemple d'accéder très facilement aux descendants d'un nœud ou de propager des informations le long de l'arborescence.

Nous donnons maintenant une idée succincte de ce que pourrait opérer une mise en œuvre parallèle de cette structure sur une machine dédiée. Cette machine doit posséder un certain nombre de processeurs indépendants, alloués chacun à un flavor particulier et liés entre eux suivant les relations d'héritage qui sont définies pour un lexique donné. Dans cette architecture la fonction **propage_fonction**() consiste simplement à lancer un ensemble de tâches simultanément sur tous les processeurs qui sont liés à un processeur donné (il n'y a pas de bouclage). Il faut aussi que tous les processeurs soient reliés à un bus de commande commun qui se contente de délivrer des instructions générales pour que chacun les exécute en parallèle. Ce mécanisme, qui est classique sur les machines parallèles (voir le n°204 de La Recherche, "Les nouveaux ordinateurs", novembre 1988), réalise directement la fonction **calc_formule**(). Dans ces conditions, on observe que les calculs proposés pour le traitement d'une hypothèse lexicale sont *proportionnels à la profondeur de l'arbre le plus long* structurant le lexique.

La maquette existant actuellement ne possède bien évidemment pas toutes ces qualités. Une trace de son exécution est fournie en annexe. Nous avons choisi un lexique réduit pour valider nos propositions, mais qui repose sur l'essentiel des traits syntaxico-sémantiques utilisés par les autres modules de notre système. Une description plus complète des éléments linguistiques utilisés est disponible dans [Romary 89]. Les deux types d'hypothèses, à savoir *apport d'information* et *demande de validation*, ont été mis en œuvre et le fonctionnement intérieur du gestionnaire est complètement transparent à un utilisateur. Bien que l'intégration complète de ce module n'ait pas été réalisée (il ne s'agit que d'une maquette), le lien direct avec une étape de vérification lexicale est immédiat et a été envisagé dans [Romary 88].

1.3.7 Conclusions et extensions.

Dans cette partie, nous avons proposé un réponse possible aux problèmes des échanges d'hypothèses lexicales dans une architecture modulaire de système de dialogue oral homme-machine. De fait, nous avons défini un module de gestion autonome qui prend en entrée des hypothèses qui lui parviennent et qui réalise correctement les deux tâches qui lui sont allouées, à savoir mettre à jour sa base de connaissances lexicales et proposer des mots probables en fonction d'un certain contexte d'analyse pour qu'ils soient validés sur le signal de parole. Pourtant, notre analyse ne semble pas complète et il est temps de faire maintenant un bilan de ces recherches.

1 POSSIBILITÉS DU GESTIONNAIRE.

Le travail effectué consistait à traiter un problème spécifique dont le cahier des charges était relativement facile à définir. Néanmoins, il apparaît que le traitement envisagé peut être généralisé en un paradigme de gestion d'informations incertaines, pour peu que l'on se limite à des domaines possédant des caractéristiques particulières : que l'univers des objets à déterminer soit fini et qu'il puisse être structuré suivant des traits spécifiques au domaine. Pour l'heure, deux applications ont été envisagées dont une est en cours de développement.

Le niveau phonétique se prête très bien à un traitement hiérarchique de l'incertain. En effet, une fois qu'un signal de parole a été segmenté (décomposé en unités cohérentes dans un spectrogramme par exemple), la détermination de la présence d'un phonème sur un segment dépend de certains traits plus ou moins nets (présence de burst, vocalisation, présence d'énergie sur certaines bandes de fréquence) déterminés par différentes

méthodes (pitch, spectrogramme, courbe d'énergie). La transposition est donc immédiate en remplaçant les mots du lexique par des phonèmes et les hypothèses que nous manipulions par des formules booléennes sur ces traits phonétiques.

Une autre application, déjà en cours de développement concerne la détermination d'objets quelconques en fonction de l'avis d'un groupe d'experts dans le cadre du projet de système à bases de connaissance ATOME ([Lâasri 88]). Là encore les objets manipulés sont souvent définis à partir de traits et l'application de notre paradigme de gestion de l'incertain est immédiat.

Nous ne détaillerons pas plus ce type d'exemples, mais on observe que toute une classe d'applications correspond à l'étude que nous avons faite et qu'il serait dommage de limiter celle-ci au simple traitement lexical. D'un autre coté nous allons voir que la souplesse de notre gestionnaire pour ce qui est du domaine d'application ne se retrouve pas quand il s'agit de compliquer le type de traitement.

2 LIMITES DU GESTIONNAIRE.

Nous avons négligé jusqu'ici de parler de la génération effective des hypothèses manipulées par notre gestionnaire. Pour chaque module (SYNSEM et DIAL principalement), doit être définie une référence commune pour les scores donnés à chaque hypothèse, mais surtout toute modification relative aux informations linguistiques qu'ils manipulent (les traits syntaxico-sémantiques notamment) doit s'accompagner d'une mise à jour de la base lexicale qui doit rester compatible avec l'ensemble du système. Cette tâche est en soi particulièrement lourde.

Le deuxième point qui n'a pas été traité concerne la gestion des intervalles temporels où un mot doit être validé sur le treillis phonétique. Cette information est très importante, car elle limite très fortement la portée d'une hypothèse informative par exemple. L'espace de décision que nous avions considéré n'est plus alors limité aux seules entrées lexicales, mais est constitué d'objets de la forme (*mot, intervalle*) où 'intervalle' est un couple de réels, en première approximation. Cet espace n'est plus fini, ce qui interdit l'usage du mode de traitement de l'incertain que nous avons développé. De plus, la gestion des informations temporelles ne se limite pas au seul niveau du mot mais est en fait très général dans un système de dialogue homme-machine, ainsi que nous le verrons aux chapitres suivants.

Enfin, il semble difficile d'étendre le gestionnaire d'hypothèses lexicales, afin qu'il agisse aussi sur l'autre catégorie d'hypothèses existant dans notre système de dialogue : les hypothèses syntaxico-sémantiques. Alors que les mots du lexique forment des unités indépendantes et en nombre fini, les structures syntaxiques par exemple se combinent les unes dans les autres et sont en nombre infini, dès que le domaine d'application du système devient un tant soit peu complexe. Pour répondre à ces contraintes, il faudrait alors développer un modèle orienté vers la notion de relation syntaxique et qui serait à nouveau inapplicable aux informations sémantiques.

A la suite des remarques précédentes, le bilan final apparaît bien sombre et le produit fini que nous avons développé n'est pas en mesure de répondre à d'autres nécessités dans notre système que celles pour lesquelles il a été conçu. Afin de poursuivre nos recherches, nous devons donc entamer une réflexion plus poussée sur ce que doit être un système de gestion de dialogues oraux homme-machine.

1.4 Critique de l'approche actuelle.

Nous venons de voir, qu'après avoir défini, dans le cadre de l'architecture d'un système de dialogue oral homme-machine, un schéma prometteur de traitement des hypothèses lexicales, nous arrivions à une impasse dès lors qu'il s'est agi d'étendre le paradigme de traitement ainsi défini aux autres catégories d'information. Il faut donc trouver les causes de cet échec et, éventuellement, y remédier, soit en apportant certaines corrections au gestionnaire lui-même, soit en reprenant à la base une réflexion sur les systèmes de dialogue. En effet, nous nous trouvons face à deux alternatives quant à la cause profonde de cette aporie :

- *le choix de la définition d'un gestionnaire d'hypothèses lexicales tel qu'il a été envisagé n'est pas le bon* et il est nécessaire de redéfinir un autre cadre de traitement de ces informations. Toutefois, alors que nous décrivions les raisons qui nous ont conduit à la définition d'un calcul de score et d'une structuration du lexique adaptée à celui-ci, nous avons vu que bon nombre de nos choix ont été imposés par la structure d'ensemble du système et des autres modules. C'est cette structure en particulier qui demande un traitement spécifique pour les différentes informations qui transitent d'un module à un autre.

- *la structure adoptée pour notre système de dialogue oral homme-machine semble donc devoir être remise en cause.* Or, cette structure est proche de celle de nombreux autres systèmes, qui semblent encourir les mêmes critiques, même si leurs concepteurs ne l'avouent pas tous ouvertement. La réflexion à mener est donc générale et l'architecture choisie par notre équipe pourra servir d'exemple pour les problèmes qui se manifestent aussi pour d'autres.

Les difficultés à gérer globalement les informations échangées dans notre système semblent provenir des grandes différences existant entre les structures internes de chacun des modules, et particulièrement de l'utilisation à chaque niveau de modèles linguistiques tout à fait différents. Ainsi nous avons un modèle hiérarchique des dialogues, une grammaire de cas pour traiter des informations sémantiques, des réseaux à nœuds procéduraux pour la syntaxe et enfin un lexique qui doit être phonétique, mais aussi qui doit être accessible à partir des informations manipulées par chacun des autres modules.

Il existe donc un problème de communication entre ces différents modules, puisque chacun gère des données différentes, et donc une nécessité impérieuse d'adapter à chaque niveau les informations en provenance d'une unité pour qu'elles puissent être comprises par une autre. On observe en particulier ce phénomène entre "SYN" et "SEM", car, malgré la réunion de ces deux unités en un seul module de traitement, une batterie de règles est nécessaire pour faire effectivement correspondre une structure syntaxique et une représentation prédicative.

Dans ce contexte, chaque module perd beaucoup de temps à transformer une forme de connaissance en une autre, alors que souvent une des données dont l'un des modules dispose peut être immédiatement utilisable, pour accroître les connaissances générales sur l'avancement du dialogue, avant même d'ailleurs qu'un énoncé soit complètement analysé. Prenons un exemple. Soit l'énoncé :

"Je suis Marocain et je voudrais obtenir une carte de séjour" .

La première partie de l'énoncé : "Je suis Marocain", permet dès sa reconnaissance de construire une représentation plus complète du locuteur, et de ce fait conduit à une prédiction sur le type de demande, que peut faire celui-ci. Ainsi, quand l'analyse syntaxique aura reconnu "je voudrais obtenir...", des expressions comme "une carte d'identité" seront délaissées au profit d'autres plus adéquates telles que "une carte de séjour". Cependant, mettre en œuvre ces mécanismes exige que de nombreux échanges soient réalisés entre SYNSEM et DIAL, et donc, à chaque fois des transformations de données (SYNSEM ne sait pas ce que c'est qu'un Marocain, DIAL, un syntagme nominal etc...). Un problème identique se pose entre LEX et SYNSEM ou DIAL, ainsi lorsque SYNSEM envoie une hypothèse sémantique relative à un lexique temporel, nous avons vu toutes les opérations que notre gestionnaire doit effectuer pour que cette information se traduise effectivement par un ensemble de mots.

Ces difficultés peuvent être resituées par rapport à certains objectifs essentiels dans le cadre de la recherche sur les dialogues oraux homme-machine ; objectifs que nous avons présentés au début du premier chapitre.

- Représenter - Chaque module de notre architecture possède sa propre représentation de l'énoncé en cours et éventuellement du contexte de dialogue. Si nous comparons cette hétérogénéité avec ce que nous cherchons à réaliser en définitive, à savoir une description cohérente de l'univers de l'application, nous sommes encore loin du but.

En particulier, on remarque combien les informations linguistiques ont un statut à part vis à vis des données sur lesquelles on raisonne effectivement.

- PRÉDIRE - Nous avons vu combien l'opération de prédiction était importante pour bien comprendre un dialogue. Nous avons ici de nombreux niveaux de prédictions qui fonctionneraient d'autant mieux s'ils pouvaient mettre leurs efforts en commun. Qu'il s'agisse de prédictions fines sur le signal, ou de prédictions abstraites du type génération de plans dans l'univers du discours.

- APPRENDRE - A chaque niveau de l'architecture, nous avons des modèles linguistiques figés qui ne permettent pas au système de s'adapter réellement à des circonstances de discours toujours en évolution. Ici, seules les informations représentées peuvent être modifiées, alors que les supports de représentation restent identiques.

Ces trois objectifs pour un système de dialogue oral homme-machine sont très ambitieux, il est vrai. En l'état actuel des recherches il ne semble pas possible de les atteindre simultanément de par l'absence de modèles généraux et de moyens de calcul suffisants. Cependant, même si nous ne pouvons proposer une solution définitive à ceux-ci, au moins pouvons-nous définir des axes de recherche permettant de conduire, à plus ou moins long terme, à la définition d'un tel modèle.

1.5 A LA RECHERCHE D'UN MODELE.

De l'étude effectuée dans ce chapitre nous avons dégagé le besoin d'échapper à une architecture de système *modulaire* beaucoup trop contraignante pour un dialogue homme-machine que nous voulons *naturel*. Cependant, seul un modèle qui intègre informations linguistiques et conceptuelles en une seule représentation serait en mesure de répondre à nos attentes. C'est pourquoi il nous est indispensable de chercher au delà du cadre restreint de l'Intelligence Artificielle que nous fréquentons tous les jours, tout en sachant que nous rejetons certaines voies qui nous paraissent dès maintenant des impasses.

Contrairement à la thèse soutenue par J. Véronis (88), nous ne pensons pas que l'étude de la communication homme-machine passe par une réduction systématique des domaines d'utilisation des systèmes jusqu'à aboutir à des situations ou l'on assure "la *transparence* de la compétence du système" à l'utilisateur (p.7). Par cette approche, le langage naturel (où la parole) se retrouve bien vite contraint à n'être qu'un langage de commande dont un utilisateur chevronné possède toutes les formules standards, mais qui nécessite une phase d'apprentissage qu'un utilisateur occasionnel refusera bien souvent. Nous verrons d'ailleurs que la notion de compétence s'oppose à ce que pourrait être un système souple et ouvert à l'utilisateur. Ceci est pour nous un objectif qui nous guidera constamment dans les développements à venir.

Nous n'ignorons pas cependant que la généralité peut être un piège pour nos recherches si elle nous empêche de résoudre effectivement les problèmes qui se posent dans le dialogue homme-machine. Cette généralité doit rester une étape de travail pendant laquelle on essaie de prendre un certain recul par rapports aux difficultés d'un domaine. C'est dans ce sens que nous avons réalisé l'étude présentée au deuxième chapitre, où le bilan des modèles de représentation linguistique et conceptuelle nous sert de référence pour situer celui que nous proposerons au troisième chapitre.

2 LES BASES D'UN MODELE COGNITIF DE REPRÉSENTATION DU DISCOURS.

2.1 INTRODUCTION.

Dans le chapitre précédent, une problématique s'est développée à partir de l'étude des échanges d'informations dans un système de dialogue oral homme-machine. Elle consiste pour nous à rechercher un modèle de représentation des connaissances qui possède certaines propriétés essentielles pour rendre un dialogue homme-machine naturel.

Avant de nous lancer dans une recherche qui pourrait s'avérer longue et fastidieuse, il est nécessaire de définir un peu mieux son objet. Dans un premier temps, nous préciserons les objectifs que nous nous sommes fixés. Puis nous verrons que de nombreux domaines de connaissance peuvent contribuer à notre quête, sous la bannière commune des sciences cognitives. Enfin, il est important de réfléchir a priori sur l'utilisation que nous pouvons faire de nos découvertes.

2.1.1 Que cherchons-nous ?

Nous cherchons un modèle, c'est à dire un outil de travail qui permette de formaliser un ensemble de phénomènes importants dans le cadre d'un dialogue oral homme-machine. Pour être utile, un tel modèle doit plus particulièrement proposer des structures de représentation, ainsi que des méthodes effectives - et, si possible, implémentables - qui nous guident pour construire celles-ci.

Nous avons insisté au chapitre précédent sur la double nécessité de pouvoir représenter:

- la langue, son organisation propre et ses contraintes et plus généralement le discours.
- l'information portée par la langue, c'est à dire l'univers du discours.

De plus, nous sommes déjà en mesure de tracer un portrait-robot du modèle recherché, par l'intermédiaire de certaines propriétés que nous aimerions qu'il vérifie. Au delà des simples possibilités de représentation, la communication homme-machine impose qu'un système de gestion de dialogue puisse effectuer des prédictions, mais surtout qu'il s'adapte au locuteur, par des facultés d'apprentissage, comme nous l'avons remarqué au paragraphe 1.4.

Enfin, rappelons que nous cherchons moins un modèle descriptif des phénomènes impliqués, qu'un modèle explicatif (et si possible opérationnel), qui fournisse des mécanismes fondamentaux permettant à tout moment de comprendre les choix effectués par le système.

2.1.2 Où cherchons-nous ?

Les lieux privilégiés de nos recherches sont bien évidemment l'informatique et, de façon plus précise, l'Intelligence Artificielle, qui cherche à simuler des comportements intelligents rencontrés habituellement chez l'homme. Cependant, l'informatique, de par les contraintes inhérentes à son support (l'ordinateur), ne permet pas toujours de prendre un recul suffisant pour analyser les phénomènes de communication homme-machine. Notre méthodologie habituelle, qui consiste à tout analyser en termes de structures de données et d'algorithmes, nous pousse en effet à travailler très vite à un niveau de détail plutôt que d'adopter une vision d'ensemble.

Or, de nombreuses autres sciences sont concernées par le langage et les problèmes de représentation de l'univers discursif ; chacune travaille pour fournir des réponses à toutes les questions qui se posent dans ces domaines. Déjà, en 1981, la conclusion de la thèse de J.M. Pierrel montrait la nécessité de poursuivre des recherches importantes relatives aux "traitements linguistiques, en liaison indispensable avec des disciplines connexes fondamentales ou appliquées : linguistique, psychologie cognitive, intelligence artificielle, etc...". Toutes ces disciplines et d'autres encore, qui entrent dans le cadre commun des SCIENCES COGNITIVES, ont pour objectif général de comprendre et de modéliser les mécanismes de la connaissance humaine. Cette convergence vers une même fin apparaît dès lors comme une chance énorme pour nos recherches à tous qui vont bénéficier d'une synergie d'idées très importante. En France, ce mouvement est entre autre représenté par l'A.R.C. (Association pour la Recherche Cognitive) dont le dernier colloque (1988) a montré qu'il existait effectivement une complémentarité des visions dans ces domaines.

Sommairement, les principales sciences qui peuvent nous être utiles dans le cadre de nos recherches sont : l'intelligence artificielle, la linguistique, la psycholinguistique, la psychologie cognitive, la philosophie, l'ergonomie, la sémiotique et les neuro-sciences. Bien qu'assez éloignées de nos préoccupations, ces dernières devront être mentionnées, car nous pensons que certains de leurs développements actuels peuvent nous être utiles.

Quant aux autres domaines, ils se justifieront d'eux-mêmes au fur et à mesure de notre exposé.

Nous n'avons pas, dans un premier temps, fait de discrimination entre langage écrit et langage parlé. Cette discrimination risquerait de nous priver, dans l'approche du langage oral, de résultats obtenus dans l'étude du langage écrit et cependant transposables. En effet, si pour les bas niveaux d'analyse, on aborde des problèmes spécifiquement liés à l'oralité (décodage acoustico-phonétique), des mécanismes fondamentaux identiques se rencontrent dans l'analyse de l'écrit et de l'oral quand il s'agit d'informations linguistiques de plus haut niveau et de représentations conceptuelles.

2.1.3 Comment cherchons-nous ?

Même si nous adoptons une attitude ouverte par rapport aux sciences qui nous sont proches, certaines adaptations seront sûrement nécessaires pour que les résultats trouvés puissent servir de base à la mise en œuvre d'un système opérationnel. Cependant, un bon modèle des mécanismes cognitifs humains peut parfaitement convenir pour un interlocuteur automatique dans un dialogue homme-machine. Sur ce point, nous sommes confortés par les ergonomes ([Amalberti 88]) qui ont montré qu'un usager se comportait de façon similaire quand il dialoguait avec un humain ou une machine, même si dans ce dernier cas les formes employées sont quelque peu simplifiées.

Soyons réalistes : nous avons peu de chance de trouver une réponse toute faite à nos besoins, puisque chaque discipline a ses objectifs propres et particuliers. Tout au plus, espérons-nous, par un large inventaire des moyens à notre disposition, mettre en évidence des tendances ou des paradigmes encore peu connus.

Nous allons articuler notre propos autour de grands sujets de recherche, plutôt que par disciplines, ce qui permettra de mieux faire ressortir les points de convergence existant entre ces dernières. Dans un premier temps, nous verrons brièvement les propositions connexionnistes pour modéliser le support de la connaissance humaine. Puis, nous aborderons le domaine de la langue et les moyens mis en œuvre pour la représenter. Enfin, nous verrons deux aspects plus spécialement cognitifs, en traitant des objets exprimés par des énoncés particuliers et de l'univers décrit par un discours complet. Nous finirons par une synthèse des grands courants qui nous paraissent importants, synthèse dans laquelle nous traiterons du statut de la logique par rapport à nos objectifs. Nous

conclurons en justifiant le choix des informations temporelles comme cadre simplifié d'étude de tous ces phénomènes.

2.2 L'APPROCHE CONNEXIONNISTE : MODÉLISER LE SUPPORT DE LA CONNAISSANCE.

2.2.1 Le connexionnisme classique.

1 ORIGINES ET PRINCIPES.

Les modèles connexionnistes reposent essentiellement sur des connaissances biologiques, puisqu'il s'agit de modéliser le fonctionnement du cerveau à partir de ses composants élémentaires : les neurones. Suivant en cela les résultats relativement récents fournis par les neurobiologistes tels que J.P. Changeux (83), les chercheurs de ce domaine essaient d'obtenir, à partir de l'assemblage de plusieurs neurones, un système possédant des propriétés d'ensemble proches de celles du cerveau, à savoir : adaptabilité, parallélisme du traitement et information répartie. Les neurones formels obtenus sont en général des modèles fortement simplifiés par rapport au neurone réel, pour lequel tous les mécanismes électriques et chimiques ne sont pas encore totalement connus.

Un exemple de neurone formel couramment rencontré dans les milieux connexionnistes ([Rumelhart 86], [Le Cun 87]) est présenté dans la figure 2.1. Il se caractérise par un certain nombre d'entrées au potentiel x_j, une sortie au potentiel y_i et une unité de traitement permettant de calculer la valeur de la sortie, étant donnée chacune des valeurs d'entrée. Par exemple, sachant que chaque entrée est pondérée par un poids w_{ji}, un calcul typique sera :

$y_i = \Theta(\sum_{j=1}^{n} w_{ji}.x_j)$, où Θ est une fonction de seuil qui peut prendre des formes très variées. Dans un réseau complètement connecté (ou auto-associatif), l'ensemble des yi est égal à l'ensemble des x_j et les bouclages ainsi créés font évoluer le réseau d'état en état à chaque calcul. Certains états stables (qui n'évoluent pas dans le temps) correspondent alors à des formes éventuellement apprises par le réseau.

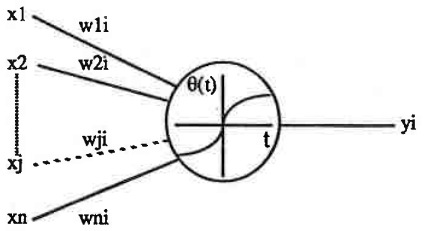


Figure 2.1 : exemple de neurone formel.

Les neurones précédents sont alors intégrés en un réseau complexe, généralement subdivisé en plusieurs couches, qui permet d'effectuer de la reconnaissance, c'est à dire une association entre une forme source et une forme cible que l'on recherche. Le fonctionnement d'un réseau connexionniste s'appuie sur deux phases principales :

- une phase d'apprentissage, au cours de laquelle on montre au réseau un ensemble de couples (source,cible) destinés à servir d'exemples pour fixer les poids w_{ji} des entrées . Cet apprentissage, généralement très coûteux en temps de calcul, repose sur différentes techniques parmi lesquelles on peut citer l'algorithme de rétropropagation utilisé notamment en France par l'équipe de F. Fogelman (voir par exemple 88).

- une phase de reconnaissance, où l'on observe le résultat en sortie du réseau après présentation d'une cible. Cette phase est particulièrement rapide et, quand elle est implémentée sur des architectures adaptées, sa complexité de calcul dépend seulement de la profondeur du réseau.

Il existe bien des variantes au modèle de base présenté ci-dessus, mais schématiquement, les mécanismes mis en œuvre sont similaires.

2 QUELQUES OBSERVATIONS.

La principale critique que l'on peut faire à cette approche résulte directement du système qu'elle essaie sommairement de copier. En effet, alors que les développements en neurobiologie sont à l'heure actuelle importants, les mécanismes qui permettent de passer du neurone à des traitements évolués, tels que ceux qui font intervenir les facultés de langage, sont loin d'être totalement connus. Il est difficilement imaginable par conséquent - avec des unités aussi grossières que les neurones formels actuels - que l'on arrive à

mettre en œuvre des facultés de raisonnement de haut niveau à court ou même à moyen terme. De fait, les systèmes réels qui sont présentés travaillent essentiellement sur des traitements perceptifs de bas niveau ("non-conscients"), importants certes, mais de peu d'utilité pour notre approche.

Un deuxième point qui peut nous faire rejeter les modèles connexionnistes dans le cadre de nos préoccupations, est leur aspect essentiellement numérique et donc, la difficulté de suivre le fonctionnement d'un réseau de manière précise. Ceci provient en particulier de l'apprentissage, qui est principalement comportemental, ce qui interdit tout contrôle extérieur pour accélérer et surtout améliorer ce processus.

Les réseaux connexionnistes ne présentent finalement que peu de propriétés intéressantes pour nous. Bien qu'ils soient en mesure d'apprendre ; les données qu'ils manipulent ne sont pas structurées puisque réparties, n'ont pas de rémanence dans le temps (tous les souvenirs du réseau sont mélangés) et surtout ne sont accessibles à l'utilisateur que sous forme de nombres totalement hermétiques.

2.2.2 Récents développements

En marge des options qui peuvent rester traditionnelles, telles que celle de F. Fogelman, ou parfois même extrémistes ([Bertille 88]), de nouvelles voies semblent pouvoir apporter au connexionnisme un renouveau et en particulier une possibilité d'aborder certains phénomènes de haut niveau qui jusqu'alors lui étaient restés inaccessibles.

Des travaux déjà assez anciens traitent de l'association directe de concepts dans des réseaux sémantiques connexionnistes [Feldman 84]. Grâce à des liens d'activation ou d'inhibition, il est possible de représenter que deux termes sont souvent associés l'un à l'autre, ou au contraire plutôt contradictoires. Ces recherches s'éloignent donc résolument des modèles biologiques pour ne garder du paradigme neuronal que le fonctionnement par connexions. Cependant, ces réseaux ne procurent pas de moyens de créer effectivement les concepts manipulés, ce qui réduit leur champ d'action à des domaines figés où les unités mises en cause sont parfaitement définies. Dès lors, l'usage de ces modèles dans un système de dialogue est difficile à envisager, puisque nous cherchons justement à assurer une grande souplesse des éléments linguistiques manipulés.

D'autres travaux dérivés des modèles connexionnistes initiaux sont en train de voir le jour au CRIN à la suite des importantes recherches du biologiste Y. Burnod pour modéliser le cerveau humain à partir d'unités plus macroscopiques que le neurone : les colonnes corticales [Burnod 87]. Le modèle proposé par Burnod permet de représenter des phénomènes cognitifs très variés, du percept jusqu'au concept, sur la base d'un formalisme unique. Il propose en particulier des solutions pour un apprentissage qui ne soit plus séparé de la reconnaissance et surtout, ce qui diffère fondamentalement des modèles connexionnistes classiques, dans lequel on puisse retrouver des structures causales par exemple, localisées au niveau de plusieurs colonnes réunies en groupements. Le fonctionnement d'un système s'appuyant sur ce modèle peut ainsi être contrôlé, et éventuellement guidé. A l'heure actuelle, l'équipe travaillant sur une implémentation effective de ce modèle ([Alexandre 88]) a essentiellement porté son attention sur des applications du type reconnaissance de formes, qui ne permettent pas de valider totalement cette approche. Néanmoins, elle reste très prometteuse, puisqu'elle cherche à répondre au même type de problèmes que les nôtres, à savoir : intégrer différents niveaux de connaissance en une même représentation structurée.

Enfin, citons ici les travaux de Béroule (85, 88a, 88b), qui a développé un modèle proche du connexionnisme classique par les éléments manipulés - des unités assimilables à des neurones - mais profondément original, puisque ces unités sont créées pendant l'apprentissage, et que ce processus s'opère en parallèle avec la reconnaissance.

Dans le cadre de son modèle, Béroule a introduit deux types de liens entre ses unités de traitement :

- des liens entre unités successives, qui représentent le contexte d'activation d'une unité particulière. Par ces liens sont exprimées toutes les informations se déroulant dans le temps.
- des liens entre une unité d'un niveau donné et des unités de niveau inférieur ou de perception. Chaque unité peut ainsi être activée par un niveau plus fin de représentation, de façon synchrone.

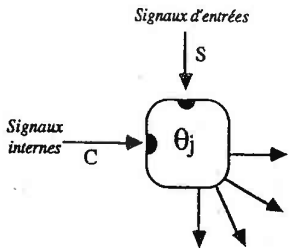


Figure 2.2 : unité de base de Béroule (88b).

Ces deux types de liens, schématisés figure 2.2, donnent une puissance de représentation toute particulière à ce modèle, visible par exemple dans l'application qui en a été faite à la reconnaissance de mots lus par un scanner. De plus, les liens contextuels traduisent un développement temporel des informations représentées, ce qui est particulièrement important pour modéliser un système de communication homme-machine intelligent, comme le fait remarquer D. Béroule (88a).

Ce modèle présente cependant le défaut d'être figé : les différents niveaux de traitement sont prédéfinis pour une application donnée. De plus, comme dans le cas des colonnes corticales, il existe encore peu de résultats susceptibles de justifier l'usage possible de ce modèle pour gérer simultanément des connaissances de bas niveau et des raisonnements sur l'univers de la tâche. Nous restons très attentifs aux développements actuels de ces travaux.

En conclusion, dans la diversité des approches dites connexionnistes, nous pouvons retenir deux grandes tendances que nous prendrons comme points de départ de notre réflexion :

- la manipulation d'objets conceptuels est beaucoup plus aisée que celle d'unités trop simplifiées dont le fonctionnement réel est souvent obscur. Ainsi, Béroule introduit différents niveaux de représentation, modélisant des mots ou des phrases.

- dès que l'on structure ces objets sur la base de relations simples et cohérentes, la puissance de représentation et de traitement du modèle correspondant s'accroit. Il semble donc plus intéressant de posséder des relations différenciées pour exprimer certains phénomènes particuliers, plutôt qu'un réseau neuronal standard n'aboutissant au même résultat qu'après de nombreux calculs.

2.3 La langue : son attrait et ses limites.

Parmi les moyens dont nous disposons pour observer les comportements cognitifs humains, la langue est assurément un outil de premier ordre. Elle semble refléter en effet bon nombre de propriétés considérées comme primordiales pour un système dit intelligent. Ainsi, elle contient dans sa structure même des schémas de raisonnement ("*si j'étais riche, j'achèterais une voiture*"), d'association d'une qualité à un objet ("*la table est verte*") ou encore des expressions prédicatives qui décrivent notre propre vision de certaines situations ("*la pierre roule*").

La langue fournit aussi de la matière à celui qui désire l'étudier, puisque de nombreuses traces écrites ou enregistrées autorisent une analyse détaillée de sa structure. Ainsi, les textes, corpus et autres sources d'informations sont de bien meilleurs matériaux de travail que l'introspection, pour étudier et donc modéliser les mécanismes cognitifs. Mais est-il possible d'arriver à un résultat par le seul truchement de la langue et de sa structure ? Nous montrerons que le problème reste posé, même si tous les courants linguistiques décrits dans cette section y apportent divers éléments de réponse.

2.3.1 Les modèles linguistiques durs.

L'étude de la langue peut s'envisager sous deux aspects radicalement opposés tant au niveau des objectifs que des méthodes. Une linguistique, qui peut passer pour traditionnelle de nos jours, préconise une autonomie complète du langage par rapport aux autres aspects de la cognition. Elle fait suite au mouvement suscité dans les années soixante par Chomsky (65) qui détache une production possible de la langue de tout contexte d'énonciation, pour ne plus analyser que les structures admises couramment par tous les locuteurs. Il oppose ainsi un modèle de la *compétence* linguistique, à ce que pourrait être un modèle de la *performance* quand le locuteur est en situation. La compétence est sensée représenter ce qu'un locuteur quelconque est en mesure de reconnaître, indépendamment du langage qu'il utilise dans ses propres productions. Ce concept est en soi difficile à appréhender, puisqu'il suppose l'existence d'une langue abstraite qui transcenderait les usages divers qui peuvent en être faits. Une confrontation avec les conséquences de cette théorie est alors nécessaire avant de l'opposer à d'autres propositions pour une approche différente de la langue et de ses mécanismes.

1 TRAITEMENT DE LA SYNTAXE.

Le modèle que propose Chomsky consiste à produire toutes les phrases *admissibles* syntaxiquement, à partir de règles permettant de décrire une *grammaire formelle* dont le fonctionnement peut alors être automatisé. Suivant le degré de complexité du langage ainsi défini, il est possible de distinguer différentes classes de grammaires, décrites dans de nombreux ouvrages tels que [Aho 72] (application à la compilation) ou plus récemment dans [Sabah 88]. Pour compléter cette théorie, Chomsky introduit l'opposition entre *structure profonde* et *structure de surface* qui reflète le lien important existant entre les deux énoncés suivants par exemple :

(1) Léon mange une pomme.
(2) Une pomme est mangée par Léon.

Alors que les phrases (1) et (2) ont une structure de surface (visible) différente, elles peuvent être déduites d'une structure profonde commune (sous-jacente) qui conduit à (1) par une *transformation* vide, et à (2) par une transformation générale d'une proposition active en une proposition passive. En adoptant ce point de vue génératif et transformationnel, on néglige le fait que les deux énoncés précédents sont bien souvent différents au niveau énonciatif, comme nous le verrons avec C. Hagège dans la section suivante, ce qui constitue une vision fortement réductrice des phénomènes linguistiques.

La grammaire générative et transformationnelle a engendré toute une série de travaux dérivés mettant en pratique ses résultats. Le traitement automatique des langues présente sur ce point un champ d'expérimentation tout à fait adapté, puisqu'il s'agit explicitement de reconstituer la structure syntaxique de textes à analyser. Dans ce domaine, M. Gross a adopté une position relativement extrême vis à vis de la syntaxe. Il nous semble important de la décrire ici, car elle est certainement significative d'un ensemble de travaux dont les limites résident justement dans les fondements même de l'approche choisie.

L'hypothèse de travail de M. Gross (68 et 77) est que la syntaxe du français peut être entièrement décrite, pour peu que l'on s'attache à mettre en évidence, pour chaque mot du lexique, l'ensemble des constructions dans lesquelles il peut intervenir. Son but est donc de constituer un dictionnaire exhaustif de la langue française qui puisse être utilisé par la suite comme une base de connaissance pour un traitement automatique de textes en français. Ainsi, pour la structure $N_0\ V_0\ V_1{}^0\ \Omega$, où N_0 est un syntagme nominal, V_0 et $V_1{}^0$ des verbes et Ω une complétive à l'infinitif, deux catégories possibles de verbes peuvent occuper la place de V_0 :

- les verbes de mouvement, comme dans "*Léon court prendre son repas*". Des restrictions de classes sont alors nécessaires puisque N_0 doit être un substantif animé et V_1 être différent d'un verbe de mouvement ainsi que de '*avoir*', '*être*', etc...
- les verbes '*oser*' et '*savoir*', auquel cas N_0 doit avoir le trait humain.

L'ensemble des descriptions ainsi obtenues pour les diverses structures, verbales par exemple, fournissent des contraintes à vérifier pour tout texte considéré comme correct lors d'une analyse automatique. Les traits éventuels qui sont introduits correspondent, selon l'esprit de cette recherche, au minimum nécessaire pour caractériser chaque structure par rapport aux autres.

Cependant, le problème devient plus complexe quand il s'agit de différencier des variantes éventuellement infimes d'un même mot ou d'un terme plus complexe. En effet, comme le remarque M. Gross (88) sur un énoncé relativement simple comme "*Le garçon entre dans le bureau, il demande le livre de compte.*", de nombreuses ambiguïtés subsistent quant à son interprétation exacte, alors même que l'analyse syntaxique a été effectuée. Il propose alors de considérer pour chaque variante d'un même terme, une entrée lexicale particulière qui, par de nouvelles contraintes structurelles locales, pourra être choisie au moment de l'analyse. Ainsi, il fournit la liste des soixante-dix entrées lexicales qu'il faudrait envisager pour le verbe manger et dont nous fournissons en exemple 2.1 un court extrait.

EXPRESSIONS SIMPLES :
Max mange sa soupe dans un bol.
Max mange à l'hôtel.
La rouille mange le fer.
Ce travail mange du temps à Max.
La barbe de Max lui mange le visage.
Les arbres mangent notre vue sur la mer.
Max mange son crayon.
Max a mangé son héritage.
Cette compagnie mange de l'argent.
Ma voiture mange beaucoup d'essence.
Max n'a jamais mangé personne.
.....
EXPRESSIONS IDIOMATIQUES :
Max a mangé du cheval.
Max mange ses mots.
Max a mangé son pain blanc.
Max a mangé le morceau.
Cela ne mange pas de pain.
Max mange sur le pouce.
Max mange à la cantine.
.....

Exemple 2.1 : Quelques formes associées à l'entrée lexicale : "manger" [Gross 88]

On constate d'emblée que le travail entrepris par l'équipe du L.A.D.L. est de très longue haleine, puisqu'il revient à couvrir toute l'étendue de la langue après l'exploration d'un maximum de textes français qui formeront ainsi un corpus supposé satisfaisant. Une fois un tel dictionnaire obtenu, son usage lors d'une analyse effective risque de demander un temps de calcul important pour retrouver la forme exacte qui convient à un lexème trouvé en contexte. Ainsi, si on se contente d'un dictionnaire de 10 000 mots pour lesquels on envisage une cinquantaine de variantes, le nombre d'entrées lexicales avoisine les 500 000 termes.

Enfin, une difficulté subsiste qui n'est pas traitée dans ce cadre : l'association aux structures de syntagmes qui ont pu être reconnues, d'objets manipulables par un système pour arriver à raisonner sur le contenu du texte lui-même. Dès lors, l'utilisation de l'importante base de connaissances actuellement disponible se limite à des travaux d'indexations de textes en vue de la création de thésaurus informatisés. Les applications à la traduction automatique [Danlos 88] se limitent quant à elles à une transformation d'une structure syntaxique de la langue source en son équivalent dans la langue cible, procédé qui reste insatisfaisant.

De façon plus générale, les grammaires formelles ont conduit à un ensemble de réalisations d'analyseurs syntaxiques dédiés à une catégorie d'applications bien précise. Le besoin d'effectuer des analyses avancées a ainsi souvent nécessité l'introduction de nombreux traits sémantiques permettant d'exprimer des contraintes supplémentaires par rapport aux seules restrictions d'accord en genre, nombre et personne. Dans ce cadre, les réseaux à nœuds procéduraux ont déjà été mentionnés. Telles qu'elles sont mises en place dans notre système de renseignements administratifs, les contraintes sémantiques ne sont pas nécessaires, puisque chaque réseau est associé, à l'intérieur d'une théorie donnée, à une construction sémantique qui introduit les restrictions nécessaires. Par contre, de nombreuses informations phonologiques ont été introduites dans les différentes implémentations des RNPs ([Pierrel 81], [Pierrel 82]) afin de tenir compte des mots outils (grammaticaux) courts dont la détection sur le signal est difficile voire parfois impossible. A cette occasion, comme cela avait été déjà remarqué pour le lexique, nous constatons la nécessité d'utiliser à un niveau donné d'analyse d'un énoncé, des informations provenant d'autres niveaux.

Une autre classe d'analyseurs développés dans le cadre de systèmes automatisés est celle des ATNs (Augmented Transition Networks, [Woods 70]), qui peuvent être vus

comme des grand-oncles américains des RNPs. Les ATNs sont des automates récursifs dont les arcs sont indexés soit par des mots ou classes de mots, soit par d'autres automates de reconnaissance lors d'une analyse plus fine d'un syntagme nominal par exemple. A chaque nœud d'un automate peuvent être associés des tests spécifiques sur l'élément en cours d'analyse ou sur les mots déjà traités, ce qui permet en particulier d'effectuer les vérifications d'accord entre les groupements qui sont construits au fur et à mesure du parcours des automates. La figure 2.3 illustre comment un sous-ensemble réduit de l'anglais peut être exprimé grâce à des ATNs.

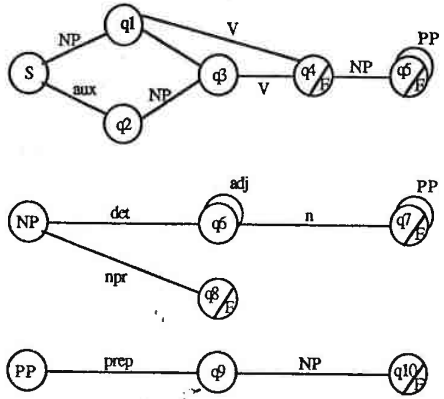


Figure 2.3 un ATN permettant d'analyser des phrases simples en anglais.[Rich 83]

2 TRAITEMENT DE LA SÉMANTIQUE.

Un traitement autonome de la sémantique semble, de prime abord, moins se justifier d'un point de vue linguistique, qu'en ce qui concerne la syntaxe précédemment évoquée. Cependant un courant particulier auquel appartient par exemple R. Martin (83), cherche à cerner le cadre d'une sémantique qui s'affranchirait de l'usage effectif d'une phrase dans un contexte, pour ne définir qu'un sens abstrait reflétant tous les sens 'immédiats' que pourrait avoir cette phrase en situation. Ainsi (p.11), "une des fonctions assignables à la théorie sémantique est la prévision *des liens de vérité* qui unissent les phrases". Préférant donc à une sémantique véri-conditionnelle couramment rencontrée en logique une sémantique ***véri-relationnelle*** qui s'appuie sur les relations d'équivalence sémantique apparaissant entre deux phrases telles que ([Martin 83]) :

p : Pierre est revenu.
q : Pierre est de retour.

A ces relations de paraphrase, dont le traitement est par ailleurs un sujet intéressant pour les sciences cognitives ([Fuchs 87]), s'ajoutent des relations d'antonymie et d'inférence qui caractérisent des couples de phrases dont le sens peut être déduit de l'une ou de l'autre. Contrairement à des relations pragmatiques, R. Martin qualifie les relations sémantiques de prévisibles et calculables, ce qui semble prometteur pour un traitement automatique. Cependant dès qu'il s'agit de déterminer la valeur de vérité d'une expression (il utilise un modèle logique de description sémantique des phrases), il se trouve face au délicat problème de rejeter tout contexte, alors que celui-ci apparaît toujours en transparence, même s'il n'est pas explicité. Ainsi, traitant de la traduction comme corpus possible pour des relations de paraphrase (note 1, p.23), il sous-entend que la transposition d'un texte d'une langue à une autre peut se concevoir directement au niveau d'une sémantique abstraite des phrases, sans référence aucune au contexte du discours dans son ensemble. Nous n'avons fait, par rapport à l'approche syntaxique, qu'un tout petit pas en avant.

Les grammaires de cas, que nous avons présentées au premier chapitre, entrent tout à fait dans ce cadre d'analyse : elles considèrent les traits sémantiques associés aux cas d'une primitive comme des indices de caractérisation d'énoncés équivalents. Cette notion d'équivalence est nettement plus souple que celle marquée par la paraphrase. Il n'en reste pas moins que la grammaire de cas permet d'exprimer le lien étroit qui existe entre les phrases suivantes :

Léon a signé la carte d'identité.
Léon a paraphé la carte d'identité.
La carte d'identité a été signée par Léon.
.....

Un autre schéma d'analyse sémantique pouvant se rapprocher de l'optique de la présente étude est l'ensemble des *grammaires systémiques*, que l'on peut trouver décrites dans [Sabah 88]. Le principe présidant à l'établissement de ces grammaires repose sur l'utilisation d'un ensemble de traits qui caractérisent des phrases, syntagmes ou mots possédant des propriétés similaires. Dans ce cadre, des systèmes qualifiant la langue en fonction de propriétés syntaxico-sémantiques comportent des informations décrites ailleurs notamment dans des grammaires de cas (cf figure 2.4). Mais ces grammaires systémiques présentent l'originalité de pouvoir exprimer l'aspect énonciatif d'une proposition sous la forme d'un système particulier (on trouvera un exemple figure 2.5).

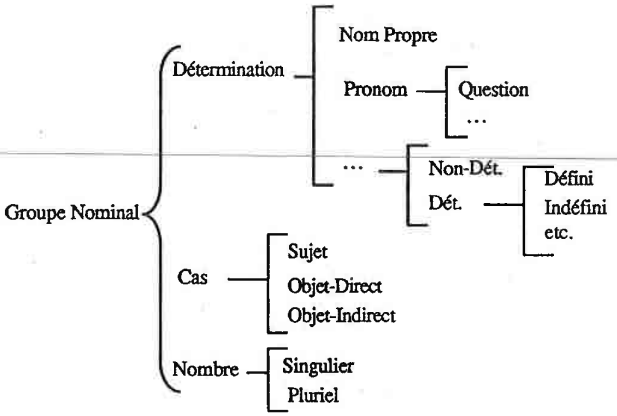


Figure 2.4 : un système partiel pour le Groupe Nominal [Sabah 88].

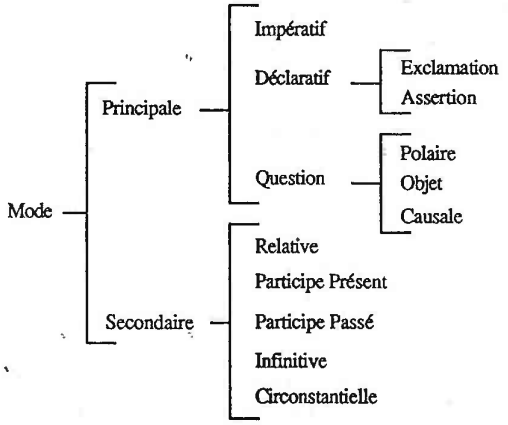


Figure 2.5 : un système partiel pour la proposition [Sabah 88].

3 CONCLUSIONS.

Nous venons ici de parcourir, de manière plus ou moins détaillée, un champ de recherche étendu qui concerne les problèmes d'analyse d'une langue naturelle, sous forme

écrite ou parlée, avec ou sans orientation délibérée vers un traitement automatique. La plupart des travaux présentés reposent sur une vision très ciblée de la langue, qui tendrait à faire penser que son objet peut être dissocié des autres domaines des sciences de l'esprit, tant au niveau syntaxique que sémantique. Cependant, l'approche choisie par les grammaires systémiques s'oriente plus nettement vers une prise en considération des phénomènes spécifiques résultant de l'insertion d'une phrase dans un contexte . Cette approche un peu différente ne pourrait-elle pas servir de base à une autre façon d'analyser les langues ? Nous reviendrons sur ce point dans la section 2.3.2.

Les théories précédentes sont caractérisées par l'utilisation systématique de catégories syntaxiques ou sémantiques figées pour décrire la langue à un niveau donné. Comme nous l'avions remarqué pour les modèles utilisés dans notre système de gestion de dialogues, cette rigueur peut rapidement devenir incompatible avec un souci de naturel pour des échanges entre l'homme et une machine.

De fait, on remarque dans tout usage de la langue une très grande souplesse de manipulation des structures sémantiques, mais aussi syntaxiques. Bien que considérées traditionnellement comme très figées, ces dernières subissent de nombreuses distorsions. Il y a bien-sûr les langages enfantins dont leurs auteurs ont peu conscience des rigueurs de la grammaire, tout en arrivant souvent à bien se faire comprendre. Sans chercher des situations aussi extrêmes, les exemples de dialogue du premier chapitre ont montré combien les énoncés parlés pouvaient être élidés et même totalement bouleversés quand l'information à transmettre est très précise en fonction du contexte. Ainsi, il n'est pas rare de rencontrer des énoncés tels que celui-là :

mon frère ? parti!

en réponse à la question :

Où est ton frère ?

La possibilité de bouleverser les structures est plus manifeste quand on touche à la sémantique. Les difficultés à déterminer des traits caractéristiques, même dans le cadre d'une application restreinte, s'amplifient encore quand on étudie les diverses possibilités de productions linguistiques couramment rencontrées. Ainsi le texte poétique est-il par définition le lieu des errances sémantiques puisque ce sont elles qui donnent à une œuvre toutes ses possibilités d'interprétation ([Eco 65]). On pourra m'objecter que la poésie n'a rien à voir avec le dialogue oral homme-machine et qu'en particulier, des textes tels que le *Finnegans Wakes* de James Joyce ne concernent pas le même langage que celui des renseignements administratifs. Cependant, le même type de distorsions se rencontre dans

le langage parlé courant, quand un locuteur cherche ses mots et remplace par inadvertance un mot par un autre de sens assez éloigné. Doit-on interdire à un usager de système ces petites imperfections et finalement l'obliger à parler *'comme une machine'* ?

2.3.2 Vers une vision plus souple des informations linguistiques.

La discussion que nous avons amorcée précédemment correspond à un débat bien plus général sur la nature du langage en rapport à la nature de la cognition, ainsi que sur les moyens à mettre en œuvre pour l'étude des mécanismes linguistiques. Nous faisons ici écho de ce débat qui fait intervenir linguistes et psychologues, car il permet de mieux situer l'opposition qui existe entre l'approche chomskienne et les travaux que nous présenterons dans les sections à venir.

1 UNE VOLONTÉ GÉNÉRALE.

Pour situer les difficultés que peut comporter l'usage d'un cadre strict pour l'étude des langues, C. Hagège reprend le thème de la traduction que nous avons déjà abordé ([Hagège 85] : p.46). Il montre combien les linguistes en général ont été tentés en analysant les analogies possibles entre différents langages, de mettre à jour des universaux de la langue, c'est à dire des sons, structures ou attitudes langagières sur lesquelles reposerait tout langage. Ces universaux seraient effectivement très intéressants pour l'objectif que nous poursuivons, puisqu'ils fourniraient la matière pour des modèles d'analyse linguistique qui seraient de plus paramétrables par la langue à étudier.

Cependant, les 'pièges et délices de la traduction' font qu'il n'est pas si facile de faire ressortir des régularités d'une langue à une autre. Chaque langue impose un cadre particulier de représentation du monde (on ne reprendra pas ici l'exemple des mots pour exprimer la neige chez les esquimaux). Pour des langues très proches telles que le français et l'anglais, des structures tout à fait naturelles dans l'une deviennent inacceptables quand elles sont traduites littéralement. Ainsi *'y aller à pied'* est difficilement traduisible par *'go there by foot'* et inversement *'walk there'* par *'marcher là'*.

La langue ne pouvant pas être définie à partir d'universaux dont l'enfant possèderait la connaissance dès son plus jeune âge, il faut admettre l'importance de l'apprentissage même si des propensions à l'activité linguistique existent dès la naissance. "La faculté de langage n'aboutit à la communication que s'il y a vie sociale". Le vieux débat entre l'inné et l'acquis est ainsi amorcé. La langue ne serait pas alors seulement un ensemble de cadres

prédéfinis dont nous apprendrions à nous servir, mais un phénomène de culture auquel les échanges sociaux nous permettraient de participer.

Si nous reprenons les termes d'un débat ayant eu lieu entre Chomsky et Piaget sur langage et apprentissage [Piaget 79], nous pouvons résumer la question ainsi :

- pour Chomsky le langage est essentiellement inné, au sens ou nous posséderions dès la naissance les structures syntaxiques que nous manipulons ultérieurement. Pour notre recherche, cela implique un modèle rigoureux (formel) du langage où ces structures sont définies une fois pour toutes sans souci de modification.
- pour Piaget, qui s'appuie en cela sur l'étude du comportement d'enfants, le langage est *acquis* au cours des années d'éducation, et bien au delà encore. Dans le même sens, A. Flieller (86) rappelle l'importance de la vie sociale dans le développement de l'intelligence. En conséquence, une théorie du langage doit prendre en compte les phénomènes d'apprentissage, ce qui correspond beaucoup mieux aux critères que nous avions fixés pour un modèle utilisable dans notre domaine d'application.

Ainsi, on retrouve constamment opposé au concept de compétence linguistique le fait que les langues se caractérisent par leur diversité [Bourdieu 82], non seulement les unes par rapport aux autres, mais au sein d'une même communauté linguistique par la liberté laissée à chaque locuteur dans l'usage qu'il en fait. Ceci est illustré par F. Latraverse (87 : p.96), qui rappelant les idées de R. Carnap, précise le statut du locuteur par rapport aux règles de la langue.

> S'il y a un sens quelconque où l'on peut parler d'un comportement correspondant aux langues formelles, il est certain que ce comportement implique les règles qui constituent ces langues. Dans le cas des langues naturelles, les règles définies par le linguiste pour rendre compte de la langue peuvent ne pas être «respectées», sans que la langue elle-même soit modifiée essentiellement ou qu'elle s'abolisse. Cela ne signifie pas qu'elles soient délibérément refusées ou bafouées par les locuteurs (ce qui supposerait que ceux-ci les connaissent, ce qui est une autre hypothèse), mais seulement que les locuteurs font autrement que ce que disent ces règles, leur comportement pouvant obéir à d'autres règles, encore inconnues. En ce sens, les régularités linguistiques sont immanentes à l'usage; si on choisit de les imputer à un système, celui-ci n'a pas d'autre privilège d'antériorité par rapport au comportement que le fait que le comportement a toujours une histoire, qu'il n'est jamais inventé tout à fait spontanément, qu'il s'effectue sur le fond de pratiques antérieures, etc.

Cependant les positions tranchées dans ce domaine ne peuvent correspondre à une réalité extrêmement complexe. L'opinion de J. Lyons (77:p.105) paraît à ce titre beaucoup moins catégorique et plus juste :

> It is perhaps impossible in the last ressort to draw a line between what is biologically and what is culturally determined, between nature and nurture.

La table 2.1 présente une synthèse des polarités qui ressortent de ce débat : la différence entre un modèle descriptif et un modèle explicatif de la langue peut y être située. En fonction des différentes remarques et critiques formulées jusqu'ici, nous chercherons en priorité du coté des modèles explicatifs qui seuls semblent être en mesure de nous apporter, en plus des possibilités de représentation, les facultés d'apprentissage et de prédiction que nous exigeons pour un système de dialogue homme-machine.

Domaines d'étude	Axes d'étude	
Linguistique	Compétence	Performance
Psychologie	Inné	Acquis
Modèle	Descriptif	Explicatif

Table 2.1 : Les deux approches possibles de la langue.

2 QUELQUES MODELES DE LA LANGUE.

A la suite du débat précédent, nous présentons des travaux qui concrétisent à leur manière l'option que nous venons de choisir pour conduire notre définition d'un modèle. Dans un premier temps, nous évoquerons les positions de C. Hagège qui se caractérisent par la définition d'une méthodologie d'étude des langues selon la théorie des trois points de vue, puis un modèle beaucoup plus détaillé sera présenté : le modèle relationnel d'Hudson.

α La théorie des trois points de vue.

Que les linguistes étudient les langues selon la phonologie, le lexique, la syntaxe ou la morphologie, ils butent sur la barrière de la phrase. Cela entraîne bien souvent les écarts que nous avons rencontrés au 2.3.1 . C. Hagège déplore cette limite en signalant combien l'insertion d'une phrase dans un texte tout entier lui donne toute sa valeur significative et il

propose une analyse suivant trois points de vue, afin d'étudier les langues "dans la réalité de leur manifestation en discours" (85 : p.207).

La théorie des trois points de vue permet de considérer une phrase suivant trois axes complémentaires et indissociables :

- le système de la langue qui donne les règles d'association des éléments linguistiques les uns par rapport aux autres. C'est le point de vue *morphosyntaxique*.
- ce qui lie les phrases avec le "monde extérieur dont elles parlent", c'est le point de vue *sémantico-référentiel*.
- le rapport avec celui qui profère la phrase ou point de vue *énonciatif-hiérarchique*. Ce dernier point concerne "l'activité du sujet dans l'exercice de la parole".

C. Hagège précise bien que ces trois manières de voir une phrase ne forment pas une suite hiérarchique qui permettrait de déduire l'un des niveaux, connaissant celui qui le précède, mais "chacun des trois points de vue apporte un éclairage d'égale importance et aucun n'est dominant". De plus "toute étude d'un seul point de vue isolé des deux autres est un artifice ignorant la réalité des liens indissolubles entre les trois". C'est une autre manière de considérer les critiques que nous avions formulées à propos des modèles prônant une autonomie de leurs traitements dans l'étendue du domaine d'étude de la langue. En particulier, cette théorie est en opposition avec la grammaire générative ([Hagège 76]) du fait de la prise en compte, dans l'analyse d'un énoncé, de la présence d'un locuteur qui crée un rapport particulier entre lui-même et ce qu'il énonce.

Afin de mieux voir comment cette méthode s'applique sur un énoncé réel, la figure 2.5 montre l'analyse possible d'une phrase simple telle que : "*Pierre chante*". On y retrouve les trois points de vue qui structurent de manières différentes les composantes de la phrase.

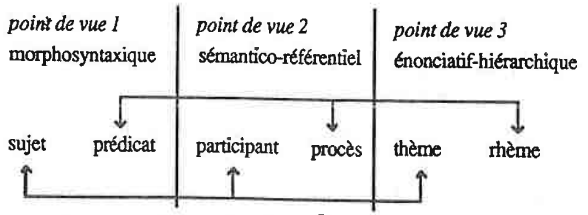


Figure 2.5 : Analyse d'un énoncé simple d'après Hagège (85).

Le cadre d'analyse proposé par Hagège n'est pas un véritable modèle de représentation. Il doit plutôt être assimilé à une méthodologie dont le but est de fournir les outils pour la définition de véritables modèles de la langue. Nous apercevons ainsi, parmi les trois points de vue, une première vision intégrée de la langue, tout en sachant que celle-ci possède malgré tout des contraintes structurelles entre ses éléments d'une part, et d'autre part un lien très étroit avec les messages qu'elle exprime.

L'aspect le plus problématique de la théorie des trois points de vue concerne le statut exact que peut avoir le deuxième point, à savoir le sémantico-référentiel. Le "monde extérieur" auquel Hagège fait référence est difficile à appréhender dès lors que l'on tient compte du locuteur par ailleurs. En adoptant une position quelque peu mentaliste, il est possible de critiquer l'association qui est faite entre une phrase et un objet ou une situation du monde extérieur, en remarquant que ce lien est culturel au même titre que le reste de la langue. Il serait alors intéressant de se demander si les deux points de vue sémantico-référentiels et énonciatif-hiérarchique ne pourraient pas être réunis suivant un seul et même axe d'analyse.

Enfin, le principal résultat des travaux de Hagège, ressortant dans sa théorie des trois points de vue, est la nécessité d'établir un véritable modèle du locuteur. C'est par celui-ci que la langue pourra être véritablement insérée dans les effets de l'énonciation, en considérant comment et pourquoi un locuteur produit un énoncé particulier :

> La seule passerelle entre sémantique et pragmatique, en acception large, dont la linguistique ait lieu d'être soucieuse, c'est le locuteur lui-même, producteur et décodeur de sens dans l'environnement social qui se constitue comme son milieu naturel. Il reste donc à l'envisager dans ce cadre. (p.233)

β Le modèle relationnel d'Hudson.

R. Hudson (84) se propose de définir un modèle général de la langue possédant les propriétés suivantes :

- il représente tous les aspects de langue, de la phonologie jusqu'à la sémantique.
- il peut être vu comme un cas particulier d'un modèle général de la cognition, au sens où les relations introduites au niveau linguistique doivent être considérées comme étant plus générales.

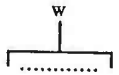
Dans la théorie de R. Hudson, la *'grammaire lexicale' (word grammar)*, le langage est vu comme un réseau d'entités reliées entre elles par des relations. Il oppose en particulier cette représentation à celle que pourrait donner une simple liste de règles et d'entrées lexicales fonctionnant selon les règles chomskiennes. Le but d'Hudson est donc de réunir en une même représentation aussi bien les lexèmes, que les structures syntaxiques ou les contraintes sémantiques. Chacun de ces objet possède les mêmes types de propriétés, et est accessible de la même manière dans son ensemble.

Les entités manipulées par Hudson peuvent donc être très variées, il peut s'agir d'éléments linguistiques tels que :

- des mots simples comme *'cours'*, *'courir'*, *'verbe au passé simple'* (classe générale), *'mot français'*, *'mot'* (classe encore plus générale).
- des parties de mots, tels que des morphèmes ou des segments phonologiques.
- des suites de mots intervenant dans des structures de co-occurrence (que nous allons définir). Celles ci peuvent être incomplètes.
- des significations de mots. Ce sont généralement des référents de mots et ils sont supposés servir de point d'entrée pour les structures cognitives.
- des éléments du contexte d'énonciations tels que le locuteur, le perlocuteur, le temps ou le lieu.

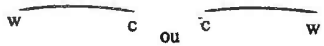
Les entités précédentes sont reliées entre elles par différentes relations permettant d'exprimer des liens représentatifs des phénomènes linguistiques et cognitifs. Les cinq sortes principales de relations introduites par Hudson sont les suivantes :

- *composition*, qui permet de lier un mot 'w' (au sens large) à ses sous-parties :



- *modèle* reliant un mot 'w' à une entité plus générale 'm' dont il est une instance.

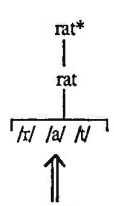
- *compagnon*.C'est la relation de co-occurence d'un mot 'w' avec un autre 'c'. Selon Hudson cette relation peut prendre des formes plus précises telles que 'sujet de', 'modifieur de', etc.



- *référent*, qui associe un mot w et une structure sémantique w*.

- *cadre énonciatif* ("Utterance-event") qui permet de lier un mot particulier avec son énonciateur, son récepteur, le temps et le lieu de l'énoncé. Hudson avoue que cette relation est en soi particulièrement complexe.

Prenons l'exemple simple du mot anglais 'rat'. Il peut être représenté localement suivant la figure 2.6.



COMMENTAIRES

référent de 'rat'.

mot étudié

décomposition phonétique de 'rat'.

classe générale regroupant toutes les voyelles.

Figure 2.6 : représentation du mot rat, adaptée d'après Hudson (84).

Un des mécanismes fondamentaux qui sous-tend cette théorie et accroît son intérêt, est la relation d'héritage. Par un souci d'économie de représentation, cette relation permet notamment à une entité lexicale d'hériter ses propriétés grammaticales d'une autre entité. Ainsi, 'chantais' hérite de toutes les propriétés de construction (transitivité, restriction sur le sujet) du verbe 'chanter'. Il est de ce fait possible de décrire les structures grammaticales de la langue, mais aussi les règles phonologiques standards (la conjugaison en particulier).

Hudson alloue à l'auditeur la tâche d'affecter chaque entité qu'il entend à un modèle dont il hérite. Après quoi, l'auditeur propage les propriétés du modèle sur l'instance, ce qui constitue le mécanisme de base de l'analyse d'un énoncé. Enfin, il peut conduire un raisonnement à partir de ces constructions, ce qui le mène éventuellement à la production d'un nouvel énoncé. Ce dernier aspect, signale Hudson, dépend plus particulièrement des

intentions de l'auditeur/locuteur dont il faudra tenir compte dans la définition des informations cognitives.

Ce dernier point fait ressortir l'aspect extrêmement mentaliste de la grammaire lexicale d'Hudson. Il affirme en effet que le langage est la propriété d'un individu - les données linguistiques sont donc subjectives - et que l'étude de la signification est inséparable de l'étude de la connaissance qu'a l'individu du monde environnant. Il affirme ainsi :

> Any theory which tries to analyse meanings without reference to meaners is doomed to circularity, prescriptivism or failure.

L'ensemble de ces informations permet de construire des représentations complexes telle que celle présentée figure 2.7, où l'on remarque que les relations initialement proposées se sont énormément complexifiées, ce qui rend l'usage du modèle assez ardu. Il garde néanmoins une très grande puissance de représentation par rapport à de nombreux phénomènes qui posent couramment des problèmes aux linguistes : homonymie, synonymie, similarité entre mots masculins et féminins, ordre temporel des mots d'une phrase, traitements sémantiques, quantificateurs, coordination et influence du contexte d'énonciation.

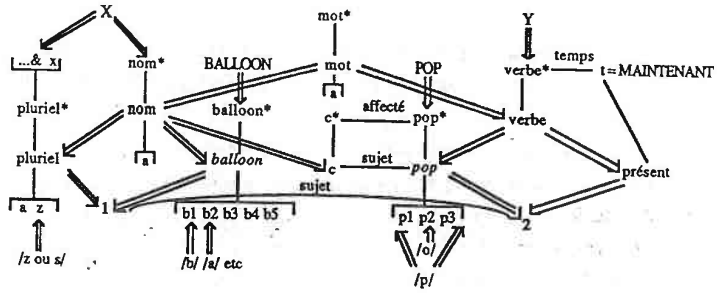


Figure 2.7 : une représentation de la phrase : "Balloons pop".

Divers points peuvent être discutés au sujet de la théorie précédente dont les ramifications sont nombreuses parmi tout ce qui touche la langue de près ou de loin. Nous ne nous attacherons ici qu'à certains aspects généraux qui sont intéressants ou critiquables.

Tout d'abord les différentes relations introduites par Hudson semblent avoir été choisies en fonction de besoins précis, et il est parfois difficile de percevoir une cohérence d'ensemble parmi celles-ci. Ainsi, pourquoi séparer les relation 'composition' et 'compagnon' qui se ressemblent beaucoup au niveau conceptuel, puisqu'exprimant les mêmes liens structuraux (suivant l'axe du temps). Pourtant, ces deux relations possèdent des caractéristiques propres qui sont tout à fait justifiables au niveau cognitif, ainsi que nous l'observerons quand nous aborderons le domaine du temps. De même la relation de référence et plus encore celle de liaison avec le contexte d'énonciation, semblent vraiment apparaître dans le but d'exprimer un aspect particulier de la langue, plus que pour donner une logique au système tout entier.

La grammaire lexicale reste aussi limitée quant aux mécanismes qui permettent à la relation d'héritage de fonctionner. R. Hudson n'introduit volontairement que peu de catégories grammaticales, signalant qu'elles apparaissent d'elles-mêmes lorsqu'une hiérarchie de structures syntaxiques est construite. Cependant, il ne montre pas comment cette construction peut être faite de manière systématique, cela impliquerait une étude supplémentaire avant d'utiliser pleinement ce modèle pour faire du traitement automatique de la langue.

Il n'en reste pas moins que le modèle proposé est extrêmement intéressant de par l'originalité de nombre de ses composantes. La comparaison avec les solutions que nous proposerons par la suite pourrait montrer son influence sur nos travaux, même si cela n'a pas toujours été conscient.

3 CONCLUSIONS.

Il existe beaucoup d'autres modèles linguistiques cherchant à intégrer les différents niveaux de représentation de la langue en un modèle unique. On pourra ainsi citer les grammaires d'unification ([Shieber 85]), qui se caractérisent par une souplesse d'analyse particulièrement appropriée au traitement du langage naturel. Néanmoins, les quelques exemples que nous avons analysés en détail présentent des orientations importantes pour une nouvelle vision de la langue :

- dans un modèle de la langue, les relations liant entre eux les objets manipulés sont bien souvent plus importantes que les objets eux-mêmes. Il ne s'agit pourtant pas de relations linguistiques, mais de relations plus fondamentales traduisant en un certain sens

des phénomènes cognitifs sous-jacents. C'est un peu en ce sens que Piaget s'oppose à Chomsky a propos des structures de la langue.

- la langue ne doit plus être systématiquement détachée des autres informations conceptuelles maîtrisées par ailleurs par tout système pensant. On en vient donc à la nécessité de modéliser dans son ensemble l'activité d'un système qui dialogue et dont les connaissances évoluent à chaque échange.

2.4 La représentation des éléments de l'univers de discours.

2.4.1 Situation du problème.

1 La langue et son message.

Malgré les incertitudes qui touchent nos connaissances du raisonnement humain, un fait a toujours été reconnu ; le langage n'est pas le seul support de la pensée. Toute une série d'autres composantes, incluant images, sensations tactiles ou odeurs interviennent pour donner des dimensions supplémentaires à nos facultés d'appréhension de notre environnement. Et s'il fallait restreindre la vision de l'univers d'un système de compréhension automatique à une seule représentation linguistique, ce serait imposer des limites arbitraires nuisibles à une qualité des échanges avec l'utilisateur.

En fait, la puissance d'expression d'un système de représentation doit inclure celle de la langue utilisée pour ne pas en limiter l'usage. De plus, elle doit être assez importante pour procurer à l'occasion une description plus fine d'un objet, d'un événement, ou d'un concept particulier. Parler d'un objet ne donne qu'une certaine vision de cet objet, en fonction du vocabulaire ou des expressions disponibles, d'où la nécessité de recourir éventuellement à des informations complémentaires.

L'étude des messages véhiculés par un texte semble être d'un niveau de difficulté bien supérieur à celui de l'analyse des structures du langage. Autant celles-ci sont accessibles directement à l'expérience, autant les significations ne présentent aucune manifestation tangible ; elles sont donc perçues de façon essentiellement intuitive. Les modèles proposés pour représenter les objets du discours vont donc s'appuyer assez peu sur des données concrètes, mais vont surtout s'attacher à traiter certains sous-problèmes particuliers, que leurs concepteurs ont pu trouver intéressants. Beaucoup de ces modèles résultent de l'analyse du langage naturel et vont, à ce titre, présenter des analogies avec des représentations sémantiques comme les grammaires de cas que nous avons déjà présentées.

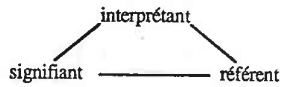
Cependant, le lien qui s'établit entre la langue et les objets qu'elle comporte est loin d'être immédiat à déterminer. Contrairement à ce que l'on pourrait croire en premier examen, le mot n'est pas le seul élément porteur de sens dans une phrase. Chaque unité linguistique, infra ou supra-lexicale peut à son tour apporter sa contribution à l'élaboration du sens d'un énoncé complet. Le morphème, en dehors des marques de conjugaison, peut

indiquer la structure de mots composés tels que *'in'* dans *'insupportable'*, *'re'* dans *'repartir'*. Le lexème, bien sûr, fait parfois directement référence à une entité plus ou moins connue : *'Paul'*, *'Marseille'*, *'Jeudi'*. Des syntagmes nominaux permettent de parler de concepts très divers : *'la voiture'*, *'la maison sur la colline'*, etc. Enfin, des propositions entières désignent des événements, au sens large du terme, faisant intervenir tous les objets ci dessus : *'le chat mange'*, *'le chat est noir'*.

Cette diversité de formes par lesquelles s'exprime la signification, s'accompagne d'une difficulté d'associer à chaque expression linguistique un objet de l'univers sensible qui nous entoure. Quelle est la signification exacte de *'le président'* dans *'le président peut dissoudre l'assemblée'*. L'expression désigne-t-elle celui qui vient d'être élu en 1988, où un président quelconque de la Vème république ? De même, dans : *'la nuit tous les chats sont gris'*, pouvons-nous réellement associer quelque chose à *'tous les chats'*, puisque nous ne les connaissons pas vraiment *tous* ? Ces difficultés nécessitent une discussion préalable qui permettra de travailler sur de meilleures bases par la suite.

2 La fin d'un mythe : le référent.

L'explication de la relation liant une expression à sa 'signification' repose au départ sur un malentendu. Parmi les vocables employés par différents auteurs, on peut choisir par exemple la représentation triangulaire utilisée par Peirce :



Indépendamment des controverses sur les termes, ce schéma fait apparaître trois entités, à savoir : le signifiant, c'est à dire le message (mot, énoncé ou autre), l'interprétant, élément difficilement définissable, servant d'intermédiaire pour saisir le référent, objet 'réel' désigné par le message.

La difficulté réside en fait dans ce dernier élément, externe au locuteur et au perlocuteur dans une communication en langue naturelle. En effet, alors qu'il semble immédiat de définir le référent de l'expression 'Mikhaïl Gorbatchev' (l'être humain connu de tous, auteur de certaines réformes en URSS etc...), cette opération devient plus complexe dans le cas de l'expression 'le président de la république' (pouvant désigner plusieurs personnes suivant le temps et l'espace de référence) et une totale gageure pour des

expressions mythiques telles que 'hydre' (pour ne pas reprendre une certaine licorne qui a déjà beaucoup servi). S'obstiner dans cette voie conduit à des impasses épistémologiques sans fin.

Ainsi, pour une expression linguistique donnée, l'affirmation possible est l'existence de "quelque chose" (sans plus de précision) présent dans l'esprit du locuteur, qui l'a amené à émettre ce message particulier. De plus, si ce message est 'compris', il conduira à l'apparition d'une autre entité dans l'esprit de l'interlocuteur, en lien avec la structure de l'expression transmise. Il ne s'agit pas de nier l'existence d'objets extérieurs à notre réalité mentale par une position idéaliste douteuse, mais de clarifier le processus de compréhension pour mieux l'analyser par la suite.

On pourrait objecter à cette position qu'en l'état actuel de nos connaissances en neuro-psychologie, nous ne savons rien de la structure effective des mécanismes de pensée. Cependant, notre but n'est en aucun cas de copier le fonctionnement 'exact' du cerveau (qui nous est inconnu), mais plutôt de proposer des modèles de représentation permettant de mettre en œuvre des systèmes dont le comportement, dans le cadre d'un dialogue avec un humain par exemple, paraisse le plus naturel possible. Evidemment, ce naturel ne pourra être obtenu qu'en s'appuyant sur certains résultats des différentes sciences de la cognition, mais ultérieurement il ne sera fait référence qu'à une modélisation particulière, non pas de l'humain, mais d'une machine vue comme élément d'une chaîne de communication.

3 Critique du béhaviorisme linguistique.

Lorsque nous cherchons à réaliser un système capable de dialoguer avec un utilisateur, quel est notre but en définitive ? Mettre en œuvre, par n'importe quel moyen, une boîte noire dont le *comportement* , tel qu'il est vu par un usager, est proche de celui d'un employé de l'administration qui répondrait à ses questions. Qu'importe le contenu de cette boîte , seules ses réactions comptent. Or il existe une approche tournée entièrement vers l'étude des comportements : le behaviorisme.

Pour le béhaviorisme les réactions humaines peuvent être déterminées de manière systématique en fonction des perceptions que la personne a de son environnement. Tout en ignorant totalement la manière dont un sujet peut avoir structuré ses percepts, le psychologue comportemental va analyser les *réponses* qu'il peut observer en fonction de certains *stimuli*. Il s'agit donc d'un rejet total de tout mentalisme, et en particulier des

notions d'idée, de croyance ou d'esprit. Au niveau linguistique, cette théorie s'est traduit par une étude des réactions associées à une expression langagière particulière, on parle alors de sémantique comportementale. On en trouve par exemple une définition dans Lyons (77) citant Watson :

> Words function in the matter of calling out responses exactly as did the objects for which the words serve as substitutes.

Cette manière de considérer le langage mérite certaines remarques, concernant plus spécialement la méthode employée, ainsi que les résultats que nous pouvons en attendre. Tout d'abord, le modèle béhavioriste est immédiatement critiquable en ce sens qu'il suppose un observateur (le psychologue) tout à fait objectif. Ceci n'est jamais vérifié comme le remarquait déjà Russel (59) :

> En omettant le fait que *lui* [l'observateur] - un organisme comme tous les autres - est en train d'observer, il donne un faux air d'objectivité au résultat de son observation. Dès lors que nous nous souvenons de la faillibilité de l'observateur, nous avons introduit le serpent dans le paradis *béhavioriste*.

Au niveau linguistique, le béhaviorisme est aussi discutable de par la vision simpliste du sens qu'il introduit en associant directement des objets à des mots de la langue. Ce point est similaire au problème posé par le référent Peircien que nous avons déjà rejeté.

Un autre caractère du béhaviorisme rend celui-ci difficilement utilisable pour définir un modèle de système intelligent qui soit implantable. Dans un premier temps, l'étude comportementale nécessite la mise à jour de couples (stimulus, réponse) qui correspondent au domaine étudié. Or, dans le cas de la langue, il est bien clair que le corpus nécessaire devient très rapidement étendu, même dans le cadre d'une application réduite. De plus, comme le fait remarquer Le Ny (79), on ne peut être sûr que le sujet a compris, ou a fourni la bonne réponse aux questions, dans le cas d'un sujet en situation de dialogue. Finalement, même si beaucoup d'observations représentatives étaient obtenues, il semble difficile d'imaginer comment elles pourraient être utilisées de manière automatique sur une machine, à moins de créer une immense table d'associations pour chaque couple d'énoncé, ce qui est assez peu envisageable.

En conclusion, les remarques précédentes nous conduisent à penser que l'approche béhavioriste ne correspond guère à nos attentes ; c'est pourquoi nous nous limiterons désormais à l'étude de modèles qui proposent une représentation interne des objets présents dans le discours. Cependant, le béhaviorisme a pu montrer l'importance de l'environnement sur le comportement d'un individu intelligent et il ne faudra pas oublier qu'un système sera plus performant s'il bénéficie, grâce à un mécanisme d'apprentissage, de l'apport de son environnement. Dans le cas contraire on arriverait à la situation d'une boîte transparente dans un univers opaque telle que présentée dans la figure 2.9 tirée de Dennett (87), c'est à dire un système manipulant ses propres unités conceptuelles et ne cherchant pas à interagir avec son environnement.

Black Box Behaviorism Black World Glass Box Perspectivalism

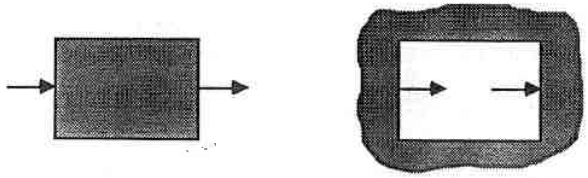


Figure 2:9 : Les excès contraires.

2.4.2 Le sens immédiat.

1 PSYCHOLOGIE COGNITIVE ET REPRÉSENTATION.

Il existe en soi assez peu de données sur les modes de représentation des informations linguistiques chez l'homme. La plupart nous sont fournies par la psychologie cognitive qui cherche à établir des liens possibles entre concepts, en partant de l'hypothèse que de tels concepts existent, sous une forme ou sous une autre. L'étude des associations entre concepts passe par la résolution de deux problèmes principaux :

- comment deux concepts peuvent être liés ?
- comment évaluer l'importance du lien entre deux concepts ?

Une expérience caractéristique des méthodes de la psychologie cognitive est celle du parcours d'image ([Kosslyn 78]). Elle consiste à présenter à des sujets une carte d'une île déserte fictive (figure 2.10), où l'on peut voir un arbre, une plage, une hutte, un rocher, un lac et un champ. Quand la carte est bien mémorisée, on demande aux sujets de focaliser leur attention sur un objet, puis sur un autre. On constate alors que le temps de

transfert d'attention entre deux objets et approximativement proportionnel à la distance entre ces deux objets sur la carte. Ceci indique alors que les liens entre concepts dans l'esprit des patients sont 'analogues' à la représentation qui leur a été fournie.

Figure 2.10 : la carte factice imaginée par Kosslyn (78).

La langue est, avec l'image, un des outils privilégiés des psycholinguistes, puisqu'elle permet d'étudier comment sont associés les concepts correspondant à des mots donnés. Suivant les cas, l'étude peut être faite soit directement sur les mots, auquel cas les fréquences d'apparition de ceux-ci prendront beaucoup d'importance, soit sur les concepts, quand on demande au sujet de manipuler ceux-ci.

La plupart des expériences effectuées reposent sur l'évaluation du temps d'accès à un mot en fonction d'un certain contexte, sur le modèle de "l'île déserte" que nous avons mentionné. La difficulté expérimentale principale qui apparaît consiste à séparer l'effet réel du mot source par rapport à un facteur fréquentiel propre au mot cible. B. Gordon (85) fait ainsi remarquer que l'accès au lexique est extrêmement dépendant de la fréquence d'apparition des mots dans la langue en général. Ce phénomène avait d'ailleurs justifié l'utilisation d'un comptage sur un corpus pour initialiser le gestionnaire d'hypothèses lexicales présenté dans notre premier chapitre. Cependant, l'effet de déclenchement (*'priming'*) d'un mot sur l'autre peut être important, du fait de contraintes syntaxiques ou sémantiques, dont les effets se combinent constamment ([Tyler 86], [Heyer 85]). Alors que la modélisation des effets syntaxiques qui concernent essentiellement des suites

de mots habituellement rencontrées se fait facilement par des théories telles que celle des cohortes de Marslen-Wilson (80), il en est autrement des déclenchements sémantiques (*'semantic priming'*) qui nécessitent l'élaboration de modèles plus poussés au niveau cognitif.

Anderson (80) parle ainsi du niveau de traitement associé à des concepts en fonction des constructions plus ou moins importantes que le sujet à pu faire autour de ces concepts. Une expérience classique dans ce domaine est celle de Bobrow et Bower (69) qui proposèrent à deux séries de sujets, soit des phrases simples sujet/verbe/objet que ceux-ci devaient mémoriser, soit simplement un sujet et un objet qu'il fallait réunir en une phrase simple. Le bien meilleur niveau de remémoration dans le second cas traduit alors un degré beaucoup plus important de traitement (*'elaborateness of processing'*), correspondant à l'insertion des objets rapportés par les énoncés dans des structures plus complexes. Une telle différence est illustrée par Anderson (p.168) par les deux réseaux de la figure 2.11, où le réseau 'a' correspond à une faible structuration conceptuelle pour la phrase "le docteur déteste l'avocat" par rapport au réseau 'b' qui intègre l'énoncé dans un ensemble de schémas le justifiant.

2 Dépendances conceptuelles et réseaux sémantiques.

Les schémas de la figure 2.11 nous permettent d'introduire très facilement un modèle de représentation des connaissances qui leur ressemble beaucoup : les réseaux sémantiques ([Quillian 68]). Le principe général consiste à lier les nœuds d'un réseau par l'intermédiaire de relations diverses telles que celles rencontrées précédemment (est_un, est_une_partie_de etc...). Ces réseaux permettent donc de représenter tout aussi bien des objets possédant des propriétés ('X' dans la figure 2.11), que des événements sur ces objets (événements 1 à 6 dans la figure 2.11).

Différents autres modes de représentation des connaissances existent dont les paradigmes sont proches des réseaux sémantiques. Frames [Minsky 75], dépendances conceptuelles et scripts [Schank 77] sont, soit plutôt adaptés à la représentation d'objets structurés, soit seulement à l'expression de relations sur ces objets. Ces représentations sont utilisées couramment en Intelligence Artificielle, en reconnaissance du langage naturel bien sûr, mais aussi dans le cadre de systèmes à bases de connaissances. Nous ne les exposerons pas toutes en détail, du fait qu'elles sont particulièrement bien connues. Il semble préférable de ne s'attacher qu'à certains points proches de nos objectifs.

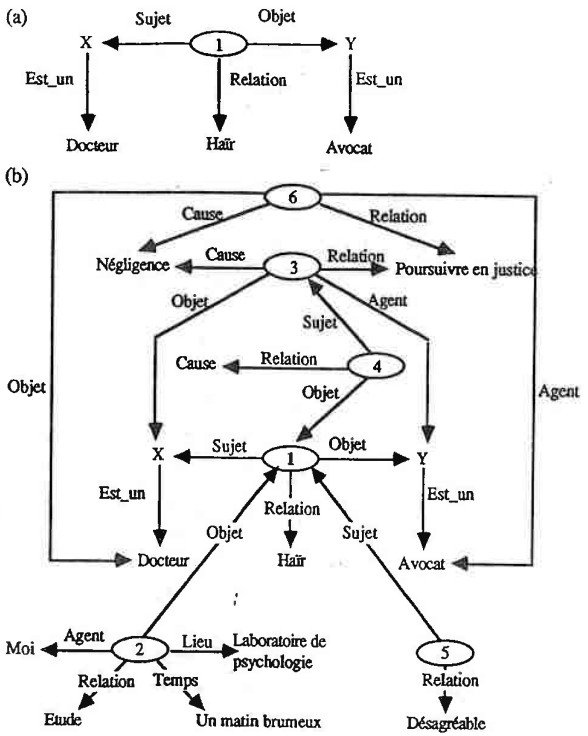


Figure 2.11 : Deux niveaux de constructions pour la phrase *Le docteur déteste l'avocat.*

Dans la plupart des systèmes de représentation de connaissances revient une relation importante : est_un (*'is_a'* en anglais). Elle permet de situer un type d'objet donné par rapport à une classe plus générale d'objets et parfois d'exprimer qu'un objet est une instance de une ou plusieurs classes. Les relations ainsi créées forment un système hiérarchique dans lequel chaque niveau hérite de toutes ou seulement certaines propriétés de ses ancêtres. Ce principe est très important en Intelligence Artificielle mais aussi pour tout ce qui touche à la cognition en général, puisqu'il donne les moyens de raisonner sur des entités abstraites plutôt que sur des instances particulières beaucoup trop nombreuses. Un domaine d'application privilégié de ce paradigme est bien sûr celui des Langages Orientés Objets tels que Smalltalk, Flavors - dont nous avons vu une application au chapitre 1 - CommonLoops, langages d'acteurs et bien d'autres encore ([Cointe 84])

La gestion d'un mécanisme d'héritage pose des problèmes dès qu'apparaissent des exceptions qui ne vérifient pas toutes les propriétés de leurs ancêtres. Divers exemples sont couramment cités et plutôt que de parler d'oiseaux (et de pingouins), la figure 2.12 présente un cas particulier traité par Touretsky (86).

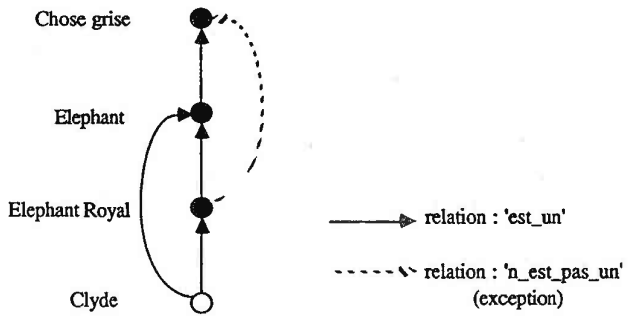


Figure 2.12 : un exemple d'héritage avec exceptions.

Si dans ce cas présenté , on cherche à effectuer toutes les inférences possibles, on arrive à une absurdité, puisque Clyde étant un éléphant, c'est une chose grise, mais puisque Clyde est un éléphant royal, ce n'est pas une chose grise. C'est là tout le problème du raisonnement non-monotone. Il nécessite la mise en œuvre de procédures particulières pour gérer ces situations de conflits dont on peut trouver un exemple formel dans [Touretzky 86].

Les travaux de D. Kayser et du L.I.P.N. ([Bonté 88], [Kayser 88]) s'inscrivent dans ce cadre puisqu'ils utilisent une logique des défauts inspirée de Reiter (80) permettant de remettre en cause un fait considéré comme vrai dans un système de raisonnement. Pourtant, ce n'est pas là l'intérêt principal de ces recherches, car D. Kayser développe avant tout un mécanisme de raisonnement à profondeur variable qui permet de décrire un objet ou une situation de manière plus ou moins fine en fonction de l'état de connaissance du système. Ce point de vue se justifie spécialement dans le domaine de la communication homme-machine où les informations échangées sont nombreuses mais souvent décrites uniquement de façon globale. D. Kayser (88) s'appuie ainsi sur le domaine du temps, dans lequel un événement est vu comme étant ponctuel ou non suivant l'angle d'analyse que l'on adopte. Alors que l'ambition du L.I.P.N. dans ce domaine est très grande, on pourra regretter que la mise en œuvre de la profondeur variable dans le système actuel

repose plus sur un artifice de calcul de cette profondeur que sur un mécanisme plus fondamental qui traduirait cette notion.

L'absence de possibilité de raisonnement à profondeur variable est un des reproches principaux qui ont pu être faits aux schémas de représentation proches des réseaux sémantiques. Cependant, on trouve certains travaux qui tentent, tout en gardant un modèle de représentation du type entités-relations, de développer des mécanismes d'abstraction au niveau des structures des réseaux qui sont manipulés. Un cas particulièrement intéressant de ce type d'approche est le système SMGC (système de manipulation de graphes conceptuels, [Moulin 88]). Partant de la théorie des graphes conceptuels de J. Sowa (84), B. Moulin propose un ensemble d'opérations visant à compléter ou réduire un graphe, instancier des variables de graphe et classer des graphes par une relation d'ordre sur ceux-ci. La figure 2.13 montre le résultat de deux opérations de jointure et de généralisation sur deux graphes donnés.

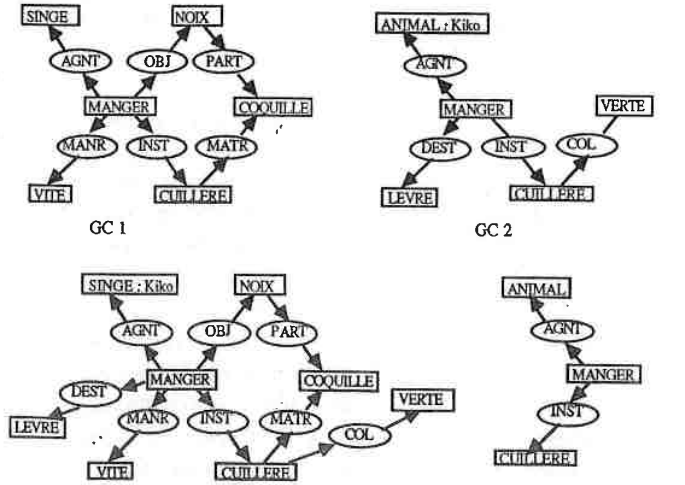


Figure 2.13 : deux opérations sur des graphes conceptuels d'après Moulin (88).

Un autre aspect intéressant des graphes conceptuels est la possibilité introduite par Sowa de composer des graphes entre eux afin de représenter plusieurs propositions intervenant dans une même phrase. Un graphe peut ainsi devenir la valeur d'un champ

d'une relation et définir de cette manière une sorte de contexte local possédant à l'occasion ses propres variables et ses propres instances. Dans la phrase "*A Québec Sam pense à Boston et pense qu'il va partir pour Boston en autobus, parce que Julie l'attends là-bas avec impatience*", on observe trois propositions dont deux forment des contextes locaux par rapport à la première comme cela est représenté dans la figure 2.14. Ces notions de contextes imbriqués sont à la base des travaux de G. Fauconnier exposés ci-dessous.

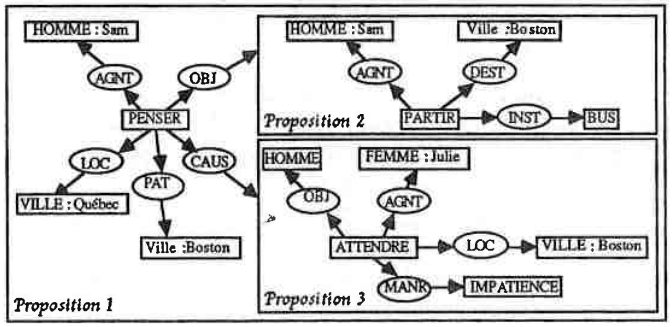


Figure 2.14 : représentation de contextes imbriqués grâce aux GCs.

2.4.3 Le monde des sous-entendus.

1 LES INCONNUES DE L'ÉNONCÉ.

Le langage ne permet d'exprimer que des objets ou des événements sous la forme de simples affirmations, mais il sert souvent de déclencheur pour relier explicitement ou implicitement un énoncé à un contexte d'analyse. On peut citer entre autres deux phénomènes fondamentaux qui s'expriment par le truchement de la langue :

- les présuppositions : tous les objets (en général) de l'univers de représentation qui sont nécessaires à la compréhension d'un énoncé. Les présuppositions ne sont jamais directement présentes dans le discours, mais doivent être néanmoins créées en fonction de la situation décrite effectivement. Par exemple, pour comprendre la phrase : "Léon a perdu sa montre", il est indispensable de la situer dans un contexte où Léon *possédait* une montre. Sans cela, la phrase précédente n'a pas de *sens*. Il vient donc que les présuppositions sont l'expression d'un contexte minimal d'interprétation d'un énoncé. Il faut alors établir des procédures pour générer ce contexte, alors que peu d'indices existent dans l'énoncé correspondant.

- un phénomène complémentaire du précédent provient de portions d'énoncés qui établissent explicitement des contextes d'analyse. Il ne s'agit plus là de présuppositions, mais plutôt de préconditions qui sont posées pour faciliter la compréhension des informations à venir. Les différentes expressions langagières correspondantes expriment alors, non pas des objets précis, mais des portions de contexte qui situent notamment la modalité par laquelle la suite de l'énoncé doit être perçue. Parmi ces modalités on peut observer des expressions de croyances ("Jean pense que...), de probabilité ("il se peut que...) mais aussi des positionnements spatio-temporels permettant de construire plus facilement les objets des énoncés.

Dans un énoncé quelconque, les marqueurs temporels occupent une place privilégiée puisqu'ils accompagnent systématiquement toute phrase simple sous la forme de temps de conjugaison. Cependant bien d'autres déictiques à valeur de temps peuvent apparaître qui se combinent pour exprimer un intervalle dans lequel tel objet ou tel événement se situe. De nombreuses analyses de ces marqueurs ont été faites ([Lyons 77], [Levinson 83]) il en ressort qu'une simple étude linguistique n'est pas suffisante pour cerner toute leur portée, mais que leur compréhension passe par une étude plus approfondie de leurs effets sur l'univers de représentation de l'auditeur. Ainsi la seule modélisation de l'énoncé et de son contenu est beaucoup trop restrictive et il faut passer à une compréhension de l'interaction entre l'énoncé et celui qui l'analyse. Nous allons voir comment les espaces mentaux sont un moyen de répondre à ces interrogations.

2 UN MODELE : LES ESPACES MENTAUX.

G. Fauconnier (85) remarque qu'à partir de certains éléments du discours, il est possible de construire des structures définissant des lieux d'interprétation des énoncés : les *espaces mentaux*. L'approche envisagée repose donc sur toute une philosophie du sens, que l'auteur résume lui-même ainsi (note 4 p.168) :

> The speaker-listener does not consider all the interpretations of a sentence and then discards the inappropriate ones. He sets up a space configuration starting from the configuration already available at that point in the discourse. There will of course be choices and strategies, but the potential of a sentence, given a previous configuration, is always far less than its general potential for all possible configurations. (A brick could theorically occupy any position in a wall, but at any stage of the actual building process, there is only one place for it to go.)

Les espaces mentaux sont alors à la fois des zones de représentation et de raisonnement qui traduisent certains phénomènes de la langue tels que les présuppositions et surtout, les problèmes d'opacité comme ceux rencontrés dans la phrase :

Selon Léon, la fille aux yeux bleus les a verts.

La représentation qu'en donne G. Fauconnier est schématisée dans la figure 2.15, où l'on observe deux mondes distincts correspondant respectivement au monde tel qu'il est vu par le locuteur (sa vision du réel) et celui de Léon tel que peut l'envisager le locuteur. Dans chacun de ces mondes existe un concept différent pour exprimer la même personne (que nous nommerons Elise), du fait que le locuteur et Léon (vu par le locuteur) ne lui assignent pas les mêmes propriétés. Un connecteur particulier permet alors de relier les deux concepts de sorte que la référence à 'a' peut servir de déclencheur pour désigner 'b'.

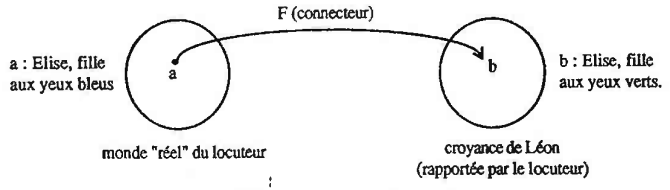


Figure 2.15 : une représentation de la phrase "Selon Léon, la fille aux yeux bleus les a verts".

Les différents mondes ne sont pas simplement décrits les uns à côté des autres au fur et à mesure de leur création, mais ils sont hiérarchisés suivant certaines règles, qui découlent parfois de marqueurs linguistiques spécifiques. La hiérarchie s'articule autour d'une relation particulière entre mondes appelée 'inclusion'. Il faut remarquer que cette relation, bien qu'elle ressemble à une relation d'accessibilité entre mondes, est très différente de ce qu'on rencontre en logique modale. En effet, elle n'introduit pas de contraintes sur les valeurs de vérité des propositions d'un espace à un autre, mais plutôt une possibilité d'accéder aux concepts d'un monde à partir d'un autre.

Un aspect de la théorie de G. Fauconnier qui peut présenter un attrait dans la continuité de notre exposé est que chaque monde ou connecteur est spécifique à un locuteur donné et résulte d'une acquisition à partir de son expérience subjective propre. Ce n'est donc pas un modèle abstrait de représentation par rapport aux agents du discours, mais bien une modélisation de l'activité de compréhension.

Enfin, on peut donner quelques indications sur la manière dont sont construits ces espaces mentaux de manière automatique. Dans de nombreux énoncés existent des marqueurs qui sont susceptibles de générer un espace dans l'univers de l'auditeur. Par exemple, l'énoncé : "Max believes that in Len's picture, the flowers are yellow", on assiste à la construction successive de deux mondes associés respectivement aux expressions "Max believes" et "in Len's picture", le dernier étant inclus (au sens de Fauconnier) dans le premier. Les expressions spatio-temporelles engendrent elles aussi des espaces mentaux, puisque justement elles situent les éléments du discours dans un contexte explicite. Ainsi, G. Fauconnier montre que dans la phrase "In 1929, *the lady with white hair* was blonde", l'espace induit par "In 1929" permet d'exprimer des propriétés particulières pour la personne qui a les cheveux blancs au moment de l'énonciation de cette phrase.

Pour conclure cette présentation de la théorie des espaces mentaux, nous pouvons résumer les différentes notions qui s'accordent avec une modélisation de l'activité de dialogue :

- les espaces mentaux introduisent une localisation de la représentation des objets transmis par le langage et donc une localisation des raisonnements les concernant. Comme dans le cas de la profondeur variable cette localisation limite le nombre de concepts manipulés à un instant donné.
- les niveaux de représentation ne sont pas figés, mais construits dynamiquement au fur et à mesure du discours. Le modèle n'en est alors que plus souple, ce qui correspond mieux à nos aspirations, sachant que nous avons déjà rejeté des théories de la langue trop classifiantes.

Le plus grand reproche que nous ferons à ce modèle est finalement de rester trop près de la langue et de ne pas procurer des constructeurs extra-linguistiques pour les espaces mentaux. Toutefois, de telles règles semblent tout à fait possibles et pourraient découler d'une extension de la théorie.

2.4.4 Conclusions.

Tout au long de cette partie nous avons étudié différentes manières d'envisager le contenu explicite ou implicite du discours, en focalisant notre attention sur certaines grandes tendances qui nous paraissent intéressantes. Ces tendances se ramènent en fait à deux aspects que nous avons continuellement retrouvés :

- *nécessité d'un mécanisme d'abstraction*. Que ce soit directement sous la forme de relation d'héritage, de profondeur variable ou par généralisation de graphes conceptuels, certaines informations communes à plusieurs objets doivent être centralisées pour simplifier leurs traitements.
- *structuration des informations*. Tout système de représentation des connaissances gagne en efficacité dès que ses informations ne sont pas mémorisées "à plat" dans la base, mais organisées les unes par rapport aux autres, comme nous l'avons vu pour les contextes imbriqués grâce aux graphes conceptuels, ou pour les espaces mentaux inclus les uns dans les autres.

Quel que soit le modèle que nous adoptons en dernière instance, il devra absolument comporter ces deux caractéristiques pour nous assurer une puissance de représentation maximale associée à une grande efficacité de fonctionnement. Avant toute chose, ces mécanismes correspondent mieux à la vision intuitive que l'on peut se faire d'un comportement naturel pour un système intelligent.

2.5 De la langue au discours : l'approche sémiotique.

2.5.1 Introduction : énonciation et compréhension.

Dans les parties 2.3 et 2.4 nous avons vu comment la langue d'un coté et l'information qu'elle véhicule de l'autre devaient être réunies en une même représentation afin de pouvoir modéliser ce qui permet de passer de l'une à l'autre et réciproquement. Cependant, les modèles rencontrés, bien que présentant certains avantages, s'attachent surtout à une énonciation abstraite, détachée bien souvent du couple locuteur-auditeur qui a beaucoup d'importance dans un dialogue. Si le locuteur produit des énoncés, ce n'est pas de manière aléatoire, mais en fonction d'un certain désir de communiquer les informations dont il dispose. Inversement, l'auditeur ne pourra assimiler un énoncé que s'il commet un acte volontaire de compréhension et dispose d'informations suffisamment complètes pour décoder celles que contient le message reçu.

L'étude de ces notions est nécessaire à la gestion convenable d'un dialogue homme-machine. En effet, un système dans cette situation de dialogue, va être tour à tour locuteur et auditeur. La recherche d'un modèle que nous nous étions fixé passe donc par une observation des réalisations en matière de production et de réception de messages linguistiques.

2.5.2 Le passage à l'acte.

A l'intérieur du couple locuteur-auditeur, c'est le locuteur qui a été le plus étudié, notamment dans le cadre des travaux tournant autour de la notion d'acte de langage. Cette vision, si elle est réduite au niveau de la phrase, se ramène à la notion d'assertion chez Stalnaker, que l'on retrouve aussi chez Le Ny quand il parle de l'acte de '*poser*'.

R. Stalnaker (84) se limite principalement à l'étude de la phrase affirmative dont les autres, selon lui, découlent par adjonction de diverses modalités. Ainsi, un acte d'assertion est, de manière essentielle, l'expression d'une *proposition*, c'est à dire d'une entité dont la valeur logique vraie ou fausse peut être déterminée. Un locuteur commettant un acte d'assertion exprime donc une certaine vérité qu'il a établie en fonction de ses croyances et de ses intentions (pour expliquer le mensonge par exemple). Ces derniers éléments forment alors le contexte, qui contient aussi les personnes à qui est adressée l'assertion, elles-mêmes accompagnées de leurs croyances et de leurs intentions. Dans le contexte ainsi défini, Stalnaker affirme que le but de toute assertion est de modifier ce

contexte, en particulier les croyances des perlocuteurs. Il y a donc une interaction constante entre le contenu d'une assertion et le contexte.

Le modèle théorique sous-jacent qui permet d'expliquer le fonctionnement de cette interaction s'appuie sur la notion de mondes possibles dérivée de la logique modale. Les mondes possibles peuvent être vus comme un ensemble d'univers d'interprétation pour lequel la sémantique d'une formule du calcul des propositions est une fonction de ces mondes dans l'ensemble des valeurs de vérité {V, F} :

$$[p] : \mathcal{M} \rightarrow \{V, F\}$$

Pour un monde M_0 donné, $[p](M_0)$ représente la valeur de vérité que prend p dans ce monde. De plus, Stalnaker introduit dans son discours une valeur de vérité intrinsèque correspondant à un monde réel qui servirait de référence. Ce dernier point est particulièrement critiquable dans notre optique, puisqu'il n'est pas évident qu'il soit possible de déterminer ce monde logique particulier et donc, a fortiori, de l'utiliser.

Les principales critiques que l'on peut faire sur ce modèle de l'acte d'énonciation concernent essentiellement le modèle logique choisi sur lequel nous reviendrons globalement un peu plus loin. Il reste que les mondes qui interviennent dans la sémantique d'un énoncé ne sont pas toujours définis très clairement. Un univers de croyance est en particulier modélisé par un monde sans souci d'une structure plus fine de ces croyances ; il s'agit alors d'une suite de propositions que le locuteur considère vraies. Cette position est critiquée par Dennett (87 : p 93) qui remarque qu'il est difficile d'imaginer que des croyances puissent être vues comme une suite de phrases au même niveau dans un espace de représentation. Enfin, la position adoptée par Stalnaker ne concerne qu'un énoncé pris en toute indépendance d'un réel *contexte* de discours formé en général d'une suite d'interventions qui s'enchaînent et se répondent. En un sens ceci est tout à fait normal, puisque dans un système logique, chaque formule peut être calculée indépendamment des autres.

Une interprétation intéressante de la position de Stalnaker peut être trouvée chez Le Ny (79), qui, à partir d'une même base logique replace l'énoncé dans son environnement. En soi, ses travaux sont déjà beaucoup plus proches de la langue puisqu'ils s'intéressent à un calcul de prédicats dans un esprit semblable à ce qui peut se rencontrer dans les grammaires de cas par exemple. Comme Stalnaker, Le Ny reconnaît l'évidence d'une norme propre à chaque locuteur, mais il parle de l'acte d'énonciation comme consistant à *poser* une *instruction verbale*. Le contexte intervient alors dans le fait qu'il situe un énoncé particulier par rapport aux autres déjà perçus, affirmant que tout énoncé est en soi un n^{ème}

énoncé. La notion de thème et de rhème dans un discours, déjà vue chez Hagège, prend alors beaucoup d'importance, car chaque thème correspond à une entité déjà définie, même implicitement.

Cette dernière notion implique de plus une mémorisation des données du discours chez le locuteur et l'auditeur. Ceci est assez nouveau parmi les théories que nous avions rencontrées jusqu'ici, dans lesquelles on avait un peu oublié que la mémoire joue un très grand rôle dans l'activité d'énonciation, mais aussi de compréhension. Un modèle utilisable dans le cadre d'un gestionnaire de dialogue homme-machine devra absolument posséder une composante gérant son activité de mémorisation, qui ne consiste pas simplement à engranger une suite d'informations, mais à les structurer et surtout à les oublier pour éviter un engorgement de la mémoire du système. Alors que le modèle logique n'est pas forcément acceptable pour une modélisation d'un dialogue naturel, la démarche de Le Ny est intéressante, car il donne certaines directions qu'il faudra suivre en définitive.

Les problèmes liés aux croyances et aux intentions du locuteur reviennent assez souvent lorsqu'on aborde les notions d'acte de langage et ce serait oublier un précurseur en la matière que de ne pas citer la *thèse intentionnelle* ('intentional stance') de D. Dennett. Celui-ci propose d'étudier tout système intelligent -au sens (très) large - sous la forme d'un système intentionnel, par opposition notamment à une étude qui pourrait porter sur la structure réelle de ce système (*'design stance'*). Cette approche permet de considérer dans un même cadre d'analyse aussi bien les systèmes biologiques que les systèmes mécaniques, ordinateurs ou autres (cf les 'two-bitsers' dans [Dennett 87] : p.290 et [Dennett 88]) et donc de transposer éventuellement les résultats de l'un sur l'autre.

Sommairement la position intentionnelle peut être résumée ainsi ([Dennett 87] : p.17) : pour un système à étudier, il faut tout d'abord le considérer comme un agent rationnel. Puis il faut chercher à définir les croyances qu'il doit avoir, en fonction de sa situation dans le monde, ainsi que ses désirs, suivant les mêmes critères. Finalement, il est possible de prédire que le système va agir en suivant ses propres buts, en fonction de ses croyances.

Bien que le modèle d'analyse soit essentiellement philosophique, et qu'il soit donc difficile de définir une procédure de détermination automatique de ces croyances, ces désirs et de leurs fonctionnements, il définit un cadre général de l'étude des croyances, qui sort un peu de l'habituel schéma logique dans lequel celles-ci sont habituellement

confinées. En particulier Dennett insiste sur la nécessité d'une structuration des croyances qu'un système intentionnel possède en remarquant que les phrases qui sont émises par un locuteur, ne représentent qu'une manière particulière de voir ses croyances, en fonction d'un choix d'énonciation et non pas l'expression directe d'une forme, qui existerait telle quelle en mémoire. Cette position n'est cependant pas incompatible avec une éventuelle utilisation du langage dans l'activité de raisonnement :

> Of course sometimes there are sentences in our heads, which is hardly surprising, considering that we are language-using creatures

2.5.3 Fabulation et Compréhension.

Un peu à l'opposé de l'activité de production de texte ou de parole, se trouve la fonction de compréhension. Si on resitue son étude dans les termes utilisés précédemment, il s'agit de reconstituer comment les croyances sont acquises et structurées au moment de l'analyse d'une portion de discours. D'après ceci, le modèle proposé par Stalnaker à partir de mondes possibles semble tout à fait adapté à cette activité, à condition de prendre la peine d'inverser son mécanisme. Chaque assertion est alors la source de l'apparition d'une proposition dans le monde des croyances de l'auditeur. Toutefois, les critiques que nous avions faites au sujet du modèle d'énonciation se retrouvent immédiatement dans ce cadre et il nous est nécessaire d'effectuer nos recherches dans d'autres directions.

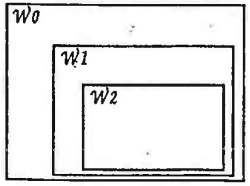
Une approche intéressante pour modéliser l'activité de compréhension est celle de U. Eco qui se place résolument dans une perspective sémiotique. Nous avons jusqu'ici assez peu parlé de sémiotique, excepté lors de l'inventaire des domaines touchant les sciences cognitives. Pourtant, ses récents développements l'ont tournée vers une étude de l'activité de communication qui a débouché sur la définition de modèles du sujet en situation, tel que celui que nous présentons ici.

Dans *Lector in Fabula* (1985), U. Eco analyse l'activité d'un lecteur devant un texte narratif et imaginaire. Il montre comment le mécanisme de compréhension repose sur un mécanisme fondamental : la Fabula, c'est à dire la génération par le lecteur de toutes les informations manquantes - non explicitées - dans le texte. Ce mécanisme est général, puisqu'à n'importe quel niveau du texte, du mot jusqu'à l'ensemble de la narration, il peut être mis en œuvre pour assurer une compréhension optimale. En particulier, la compréhension d'un simple mot ne se fait pas par sélection d'une valeur sémantique

parmi d'autres, mais en fonction des fabulations que le texte a engendrées, et dans lesquelles le mot peut s'insérer (ce point est un aspect essentiel de [Eco 72]: la structure absente, voir aussi la note 4 p.168 de [Fauconnier 85] et [Lyons 77] p.46).

Le modèle effectif proposé par U. Eco pour qu'un lecteur puisse effectivement produire cette activité, repose lui aussi sur la notion de mondes possibles, mais de façon fort éloignée de l'acception logique de Stalnaker. En effet, Eco s'intéresse beaucoup plus à la structure que peuvent adopter ces mondes, plutôt qu'aux valeurs de vérité de certaines propositions. En effet contrairement aux mondes possibles de Stalnaker , les mondes de U. Eco ne sont pas des entités abstraites qui pourraient être fournies par un observateur 'en blouse blanche', mais il s'agit d'un réelle modélisation de l'univers cognitif du lecteur, seul devant son texte. C'est en soi un pas important pour nous, puisque notre système de dialogue est en quelque sorte dans cette situation, face aux énoncés de l'usager.

Un monde particulier est un ensemble d'informations structurées connues du lecteur, contenant éventuellement d'autres sous-mondes. Par exemple, à ce que Stalnaker appelle monde réel, on peut faire correspondre un niveau supérieur (peut-être maximal) correspondant aux connaissances brutes ou perceptives du sujet. Un sous-monde peut alors être l'univers de connaissance formant le contenu d'un texte qui, de nouveau, génère d'autres mondes, correspondant par exemple à l'univers de croyance des personnages du récit, tels qu'ils sont perçus par le lecteur. Toute une structure complexe d'enchâssement peut ainsi exister permettant d'expliquer en particulier pourquoi un texte est plus ou moins bien compris en fonction de présuppositions qui sont faites quant à la structure des mondes. Dans le cas d'un dialogue, il est facile de transposer ce modèle sous la forme d'un moindre nombre de mondes présentés sous une forme simplifiée dans la figure 2.16.



W_0 : monde du système

W_1 : monde du locuteur vu par le système

W_2 : ce que le système pense que le locuteur sait à son sujet

Figure 2.16 : modélisation du monde du système en situation de dialogue. Adapté d'après [Eco 85].

Ce modèle est en quelque sorte un complément de l'approche de Fauconnier à un niveau plus élevé d'abstraction dans le discours. Il semble qu'en comparant les deux approches, on puisse effectivement mettre en correspondance les mécanismes proposés par chacun. Ainsi, Eco introduit la notion d'accessibilité entre mondes, qui permet de définir quelles sont les connaissances d'un monde donné utilisables dans un sous-monde, problème qui est très proche des phénomènes d'opacité dans les espaces mentaux.

2.5.4 Conclusion.

Dans cette partie nous sommes progressivement passé d'une vision horizontale des informations transmises dans le discours à une vision beaucoup plus structurée dans laquelle le locuteur (lecteur, auditeur, etc...) occupe une place beaucoup plus importante. Bien qu'aucun modèle ne se soit détaché de façon probante pour être utilisé dans le cadre de nos recherches, des orientations plus qu'intéressantes sont apparues qui font écho aux autres théories rencontrées dans les sections précédentes. Une synthèse de tous les aspects rencontrés s'impose maintenant avant de proposer notre propre méthode d'analyse.

2.6 Des oublis et des tendances.

2.6.1 Et la logique ?

1. Introduction.

Jusqu'à maintenant, nous n'avons que peu abordé les modèle logiques de représentation des informations linguistiques. Ils représentent pourtant une très grande quantité de travaux qui traitent de nombreux sujets envisagés dans les chapitres précédents. Cette mise à l'écart résulte en fait du désir que nous avons d'explorer de nouvelles voies. Cependant, les modèles logiques sont à la base de bon nombre de réalisations en Intelligence Artificielle en général. Cette popularité résulte essentiellement de l'ancienneté des recherches qui, sans remonter jusqu'à Aristote et l'*Organon*, ont été très développées au XVIIIème siècle (Kant) et surtout depuis les années trente autour des problèmes de logique mathématique. Ces travaux féconds ont conduits à des constructions théoriques solides et autonomes particulièrement attrayantes quand on recherche un modèle pour un système de raisonnement.

Il est important de distinguer deux usages différents dès que l'on parle de logique (ou des logiques) en général pour apprécier les avantages et les défauts de celle-ci. Il y a d'un coté une théorie de la logique qui correspond à une réflexion sur des systèmes *formels* bien définis à partir d'une syntaxe et d'une sémantique, et possédant des propriétés déductives très fortes telles que la *cohérence* ou la *complétude*. Du fait de ces propriétés, on est assuré de la qualité des déductions obtenues grâce à ces systèmes, une fois posés les outils de base. D'autre part, nous devons considérer les applications qui sont faites des systèmes théoriques précédents en vue de faciliter l'étude de phénomènes "naturels" et la simulation de ceux-ci. Nous résumons les applications qui nous concernent dans la table 2.2 où pour les deux activités de représentation et de raisonnement, nous avons mis en évidence les domaines d'activités plus spécifiquement concernés. Les flèches en grisé indiquent d'ailleurs que cette séparation n'est finalement qu'une tendance.

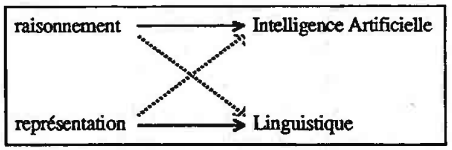


Table 2.2 : Utilisation de la logique dans différents domaines.

2. Cas du raisonnement.

Les propriétés que nous avons mentionnées à propos de la logique sont très utiles pour un système reposant sur ce type de modèle et désirant aboutir à un résultat sûr. Ainsi, les moteurs d'inférence de nombreux systèmes experts reposent essentiellement sur la logique des prédicats du premier ordre à quelques variantes près. Le but de ces programmes n'est pas alors de copier le raisonnement humain mais plutôt de compléter par leurs performances les connaissances apportées par les experts. L'exemple le plus couramment cité est ainsi Prospector ([Hart 78]) dont la mise en œuvre a permis la découverte de minerais de cuivre et d'uranium. On peut comparer l'esprit de ces approches à celui de l'utilisation de théories avancées des mathématiques pour effectuer de nombreux calculs numériques pour des problèmes de mécanique des fluides. Dans les deux cas il s'agit d'obtenir des résultats relativement précis et rapides qu'un homme ne serait jamais en mesure de fournir.

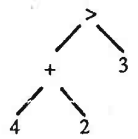
Pourtant Daniel Kayser (1987) discute l'efficacité réelle des modèles logiques et particulièrement quand il s'agit de les appliquer au raisonnement de sens commun, c'est à dire pour les activités telles que la perception, la motricité ou la compréhension du langage dans lesquelles nous excellons tous les jours. Il affirme ainsi (p.83) : "Les premières tentatives d'informatisation montrent que les raisonnements qu'il faut mener ne ressemblent guère à ceux des logiciens". La difficulté consiste alors à ne pas assimiler le raisonnement logique au raisonnement humain et d'expliciter ainsi clairement les différences qui existent entre les deux. Sans faire d'étude exhaustive à ce niveau, nous pouvons aborder rapidement la notion de valeur de vérité qui, si elle est fondamentale au niveau de la théorie logique, apporte peu à un système de raisonnement manipulant beaucoup d'informations. Ainsi, un utilisateur de système expert a rarement l'occasion de poser des questions auxquelles le système répondrait par oui ou par non. Il cherche en effet beaucoup plus à obtenir de son système de l'information et donc des réponses construites, ce qui ne correspond pas forcément à l'esprit de la logique. Prenons deux exemples parmi beaucoup d'autres qui illustrent ce phénomène. Le langage Prolog, qui

est une application directe du calcul des prédicats, ne se contente pas de donner une réponse positive ou négative à une question qui lui est posée. Il fournit les affectations de variables qui lui ont permis de répondre, donc des informations supplémentaires. Pour parler un peu de linguistique, on peut observer combien Kamp (84) a essayé de relativiser la notion de valeur de vérité au niveau des parties du discours, pour finalement attacher beaucoup plus d'importance à l'intégration de nouvelles connaissances apportées par un énoncé à un volume de données déjà existantes.

Finalement, nous avons progressivement déplacé notre propos du problème du raisonnement à celui de l'information et donc de la représentation de celle-ci. Alors que dans des applications telles que la modélisation de programmes informatiques (travaux de Hoare, Manna, Pneuli, Wolper...), la logique est encore très utilisée sous sa forme déductive, il apparaît qu'en Intelligence Artificielle et surtout en linguistique, sa capacité à représenter soit bien plus importante. Cependant il n'est pas certain qu'elle soit conçue pour cela et nous allons essayer de voir pourquoi.

3. Cas de la représentation.

L'application de modèles logiques à la représentation des informations portées par le langage résulte d'une similarité qui existe effectivement entre certaines structures sémantiques de phrases affirmatives et une représentation sous la forme de prédicats par exemple, comme nous l'avons déjà observé dans les grammaires de cas. Cependant, un des reproches que l'on peut faire à de telles applications est d'entretenir une ambiguïté constante au sujet des termes 'syntaxe' et 'sémantique', quand ils concernent une langue naturelle ou le système formel que l'on manipule. Dans ce dernier cas, la syntaxe est un ensemble de règles formelles qui définissent la structure que *peuvent* et *doivent* prendre les formules manipulées par le système logique et la sémantique donne une manière de calculer ces formules pour éventuellement leur donner une valeur de vérité. Par exemple, la formule suivante représentée sous forme d'arbre :



donne la valeur vraie quand elle est interprétée dans l'ensemble des entiers naturels $\mathcal{N}$ de façon canonique, de sorte que '+' est associé à l'addition dans $\mathcal{N}$ et '>' à la relation d'ordre stricte sur $\mathcal{N}$. Sommairement, tout système logique repose sur des bases théoriques similaires qui en font des outils rigoureux mais contraignants.

A l'opposé, les langues naturelles ne sont pas du tout des ensembles de structures figées dont la sémantique (entité qui à ce stade de l'exposé n'a toujours pas été définie et n'est pas près de l'être) se calcule de manière immédiate dès que l'on dispose de la structure syntaxique. Or, l'ambiguïté n'est pas levée, puisque l'on rencontre des travaux cherchant justement à effectuer de tels calculs ([Nef 84] et particulièrement [Nef 88]:p.83) dont les plus connus peuvent être ceux de Montague, descendant directement de la lignée générativiste. De même, comme souvent remarqué, la sémantique des langues naturelles n'est pas calculable, du fait justement qu'elle est le reflet d'un monde (ou d'espaces de cognition) complexe.

La langue s'adapte mal aux structures logiques ; inversement, beaucoup ont essayé d'adapter la logique aux langues naturelles afin de traiter certains problèmes spécifiques. Un cas exemplaire d'une telle adaptation est l'utilisation des logiques modales sous la forme de mondes possibles comme ceux de Stalnaker. Les modifications de ce type apportées au cadre initial de la logique sont nombreuses et peuvent même mener assez loin. Nous ne citerons comme exemple que la logique développée par Shoham (88b), en vue de traiter certains problèmes temporels et qui est une «logique temporelle modale et non-monotone». On obtient ainsi des systèmes patchwork dont la cohérence d'ensemble devient assez difficile a évaluer. De plus, comme l'a montré D. Kayser (87), de nombreux aspects différents (typicalité, analogie, évolutivité, apprentissage etc...) doivent être pris en compte pour représenter des informations et raisonner sur celles-ci. Il semble assez difficile d'envisager que des assemblages de théories différentes mènent à un résultat satisfaisant par rapport à ces problèmes.

4. Bilan.

Plutôt qu'un rejet total des approches logiques, nous rejetons finalement un usage excessif de celles-ci pour prendre en compte des phénomènes pour lesquels elles ne sont pas toujours adaptées. Il semble qu'il soit beaucoup plus important, au moins dans le cadre d'une recherche fondamentale, de réfléchir sur les problèmes du dialogue homme-machine (et donc du langage) dans leur ensemble sans chercher désespérément à adapter certains modèles dont l'objectif n'était souvent que de traiter certaines sous-parties de ces problèmes. Par rapport à la logique, il existe donc deux axes possible d'étude qui rejoignent la distinction que nous avions faite au début de cette partie :

- d'un coté il est nécessaire de réfléchir à des modèles logiques nouveaux, plus aptes à traiter la modélisation du raisonnement de sens commun, mais dans un cadre théorique bien défini et non par accumulation d'anciens modèles. Les travaux plus ou moins récents

en logique naturelle et surtout en logique linéaire ([Girard 87]) montrent que c'est une voie possible.

- d'autre part, pour les chercheurs en Intelligence Artificielle, qui ne sont pas nécessairement des théoriciens de la logique, il peut être intéressant d'aborder d'autres horizons qui seraient eux-aussi en mesure d'apporter des réponses aux problèmes qui se posent. C'est en ce sens que nous avons réalisé l'étude présentée dans ce chapitre, et que nous proposons un modèle au troisième chapitre.

2.6.2 Conclusion : de l'homme neuronal à l'homme dialogal.

Au cours de ce chapitre, diverses orientations scientifiques sont apparues pour traiter de la langue en tant que support de communication dans des cadres parfois beaucoup plus généraux que le dialogue homme-machine. Cependant les analyses de détails que nous avons effectuées nous donnent de bonnes raisons de rester attachés à quelques grands choix primordiaux.

Tout d'abord, nous avons trouvé ce que nous cherchons à modéliser en définitive : le mode de fonctionnement cognitif d'un système en situation de dialogue. Ceci comporte d'un coté l'usage de la langue, qui devient une des composantes de la cognition, mais aussi la représentation et l'usage qui sont fait du contenu de celle-ci. De plus, la représentation des phénomènes considérés, pour être efficiente, doit posséder une structure reposant sur des mécanismes qui semblent fondamentaux au niveau cognitif. Parmi ceux-ci nous avons rencontré en particulier les problèmes d'abstraction et d'univers emboités. Enfin, il ne suffit pas de posséder des représentations des objets de l'univers, mais il faut établir l'usage qui en est fait, surtout lors de l'analyse d'un énoncé. C'est tout le sens de la Fabula de U. Eco qui montre que beaucoup plus de représentations sont construites à partir d'un texte que les seules informations exprimées explicitement par celui-ci.

Nous nous détacherons donc d'une étude rendant transparent le fonctionnement d'un système intelligent, pour chercher à comprendre ce que peuvent être les mécanismes sous-jacents à des comportement de communication. Plutôt qu'un système neuronal incompréhensible, notre objectif le plus lointain est la conception d'un système dialogal ouvert.

2.7 Temps et cognition.

2.7.1 Justification d'une étude réduite au temps.

La définition d'un modèle du locuteur-auditeur en situation de dialogue tel que nous venons de l'envisager est une tâche ambitieuse qui correspond avant tout à un objectif à long terme. En effet, dans ce sens, peu d'avancées sont visibles jusqu'ici et il serait malhonnête de prétendre aboutir à un modèle complet et autonome en peu de temps. Il faut cependant trouver un moyen d'aborder la définition d'un tel modèle sans pour autant se perdre dans l'étendue des problèmes à traiter dans le cadre d'un dialogue homme-machine.

Le choix que nous avons fait consiste à analyser un sous-problème particulier touchant à tous les aspects de la langue et de sa représentation : le temps. En effet, la mise en place d'un système de représentation des connaissances, destiné à interpréter des énoncés oraux en langage naturel et agir simultanément sur son environnement (par une réponse ou une action physique), impose la prise en compte du temps à plusieurs niveaux.

Tout d'abord, un énoncé oral s'appuie sur un signal temporel dont l'analyse passe par un découpage en unités linguistiques qui se succèdent dans le temps. Ceci est vrai à tous les niveaux de la langue, dans le dialogue par exemple, dont les différentes phases peuvent être vues comme des événements linguistiques à forte composante temporelle. Ainsi, la phrase "le chat dort", prononcée dans une situation d'interlocution définie, se compose de trois événements "le", "chat" et "dort" qui apparaissent successivement pour l'auditeur. L'analyse des structures de la langue peut donc passer par une analyse temporelle de celle-ci.

Parmi les objets véhiculés par le langage, de nombreuses références sont faites à des événements ou simplement à des 'intervalles' de temps de la vie courante. Il y a tout d'abord les temps de conjugaison qui permettent de situer l'action de l'énoncé par rapport à celui-ci : "la maison était rouge", "la maison est blanche", "la maison sera verte". Par ailleurs, des locutions permettent de situer cette action plus précisément : "dans la nuit", "hier", "dimanche matin". Enfin, l'énoncé lui-même dans sa globalité présente une action particulière qui nécessairement se déroule dans le temps de façon plus ou moins importante : "un coup de feu claqua", "j'ai fermé la porte", "la terre tourne sur elle-même"...

Enfin, l'ensemble des objets temporels cités ci-dessus ne sont pas seulement des constructions dans un espace de représentation, mais servent de base à tous les raisonnements relatifs à l'univers du discours, notamment pour l'étude des causes et effets d'un événement précis.

Ainsi, tous les problèmes qui ont été envisagés précédemment par les linguistes, psychologues, informaticiens, philosophes ou sémioticiens peuvent être reformulés et traités sur la base d'une réflexion sur le temps et de sa représentation. Cependant, il ne faut pas retomber dans les défauts des modèles que nous avons rencontrés, qui cherchent à représenter de manière autonome certains points très précis sans se préoccuper des autres. Les objets temporels du discours doivent donc être considérés dans leur ensemble, sous la forme d'une représentation unifiée.

Le choix de cet axe de recherche semble transcender l'ensemble des problèmes de traitement et de représentation qu'il faut aborder dans l'étude de la langue. De plus il en donne une vision plus simple qui permet de faire des propositions de modèle sans se noyer dans les détails d'une analyse exhaustive. Par la définition d'un modèle de représentation des informations temporelles, nous pourrons peut-être entrevoir, à plus long terme, les fondements d'un modèle intégré pour un système de dialogue oral homme-machine.

2.7.2 Les modèles de représentation du temps.

Afin de mieux situer le modèle de représentation des informations temporelles que nous proposons, nous allons évoquer les principales méthodes existantes en intelligence artificielle et ailleurs traitant du temps et des phénomènes connexes.

1 LE TEMPS CACHÉ.

Le temps a longtemps été ignoré dans de nombreux travaux où il semblait naturellement intervenir. Ainsi la plupart des programmes de génération de plans traditionnels (par exemple STRIPS dans [Nilsson 80] p.298) s'appuient sur un ensemble d'états et d'opérateurs permettant de passer d'un état à un autre, sans tenir compte du temps de réalisation d'un tel opérateur, ni de la durée d'un état entre deux applications d'opérateurs. Les systèmes obtenus ne sont alors que des manipulateurs d'expressions formelles, bien loin des réalités d'un monde en continuelle mouvance. La représentation du temps est en définitive implicite dans ces systèmes et peut être schématisée comme

dans la figure 2.15 par un arbre des états possibles découlant d'un état donné E0. Il s'agit là d'un temps arborescent, sans passé et donc uniquement prospectif.

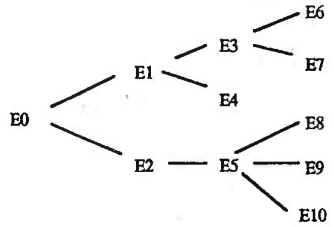


Figure 2.17 : Modèle du temps dans un générateur de plans.

2 LE TEMPS EXPLICITE.

α Le temps physique.

Le temps est une notion difficile à appréhender en elle-même et la vision que nous en avons dépend souvent de l'enseignement que nous avons reçu à ce sujet. Les sciences physiques présentent par exemple un temps uniforme modélisé par l'ensemble des réels de $-\infty$ à $+\infty$ et passant par une origine t_0, qui peut être le début de l'expérience réalisée. Or, les nombres réels se caractérisent par certaines propriétés très fortes, que le temps se doit de vérifier par définition. En particulier, entre deux nombres différents il existe toujours une infinité continue de réels, ce qui donne à ce modèle du temps une granularité aussi fine que possible et perpétuellement à disposition. Pourtant, dans la plupart des expériences de physique, on ne se sert jamais d'une telle précision de description des phénomènes. De plus, les possibilités cognitives de l'homme ne lui permettent pas d'avoir conscience d'une infinité d'instants correspondant à une infinité de situations.

β Le temps psychologique.

Le modèle physique est donc difficilement applicable à un dialogue homme-machine car l'on désire éviter de noyer l'usager dans un océan de détails qu'il ne pourrait assimiler. Il faut donc s'orienter vers une vision plus cognitive du temps, comme le prônent des psychologues tels que José Morais dans ([Alegria 83] : p.150) :

> Nous ne percevons pas le temps, nous percevons des événements qui ont une certaine durée, qui se succèdent dans un certain ordre. Et nous ne percevons pas l'espace, nous percevons des objets qui ont une certaine étendue et qui sont dans certains rapports de position, d'orientation entre eux et par rapport à nous. L'espace et le temps sont des modalités générales de toutes nos perceptions : même percevoir est une activité qui s'effectue dans l'espace et dans le temps.

Ce temps psychologique est confirmé par Piaget (46) qui insiste sur plusieurs points fondamentaux :

- *le temps est perçu de manière relative.* Les différents événements perçus par un sujet sont repérés les uns par rapport aux autres et non pas dans l'absolu par référence à un calendrier dont on aurait une conscience en soi.
- l'apprentissage de la notion de temps passe par une faculté de raisonner à *profondeur variable*, pour ne pas se perdre dans les détails d'un ensemble d'événements qui se succèdent (p.22). Ce changement de niveau est d'ailleurs indispensable pour avoir une vision d'ensemble d'une suite d'événements : "une succession de perceptions ne constitue pas à elle seule une perception de la succession, ni *a fortiori* une compréhension de la succession".
- la perception de la causalité est dérivée de celle du temps. La causation n'est alors qu'une manière de voir un schéma couramment rencontré de succession d'événements.

Ces données sont fondamentales pour l'établissement d'un modèle de représentation du temps dans un dialogue homme-machine. Elles situent le type d'opérations qu'il faut mettre en œuvre pour que le locuteur se sente en situation naturelle quand il raisonne avec le système sur des informations temporelles.

γ Le temps linguistique.

Le temps est un cadre d'étude intéressant pour les linguistes, puisqu'il intervient à différents niveaux d'analyse de la langue. C'est pourquoi les travaux dans ce domaine sont nombreux et il serait difficile d'en faire un inventaire exhaustif . Tout d'abord, on constate que les expressions linguistiques liées au temps se prêtent difficilement à une classification, en particulier sur des critères syntaxiques, indépendamment de la sémantique de ces expressions ([Borillo 86]). Ce point est déjà satisfaisant pour notre étude réduite, puisqu'elle doit prendre en compte la sémantique du dialogue.

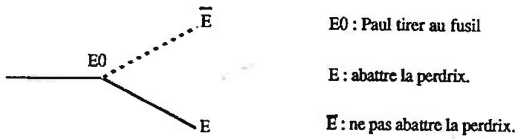
Une étude approfondie, liant les expressions à nature temporelle et leur représentation éventuelle sous forme conceptuelle est donc nécessaire. Deux travaux sont particulièrement intéressants dans cette voie.

Hornstein (77) à la suite de Reichenbach (47) s'est intéressé à la représentation du temps de conjugaison dans le discours, et en a proposé une modélisation faisant explicitement référence à l'acte d'énonciation. Il a ainsi montré qu'avec trois marqueurs temporels particuliers, à savoir S (l'instant de l'énonciation), R (une référence temporelle) et E (le temps de l'action décrite dans l'énoncé) il était possible de décrire la plupart des phrases complexes faites d'une proposition principale, de subordonnées de locutions temporelles (des déictiques notamment). Afin de localiser ces trois points, il introduit deux relations, une qui situe un des points avant l'autre, et l'autre, indiquant une proximité ordonnée entre deux points. Ces deux relations sont symbolisées respectivement par '_' et ','. A titre d'exemple, le prétérit peut être représenté par la séquence : 'E,R_S'.

Deux critiques principales peuvent être faites sur ce modèle de représentation du temps ('tense'). Premièrement, il est difficile de donner une signification claire au point de référence en tant qu'objet temporel. En fait, il apparaît avant tout qu'il doit être utilisé comme un outil de calcul plutôt que comme un réel objet de l'univers. La deuxième remarque concerne la relation de combinaison entre deux formules temporelles qui conduit à des absurdités comme le fait remarquer Yip (85), du fait qu'elle n'est pas commutative, et qu'elle exprime une sorte de précédence entre deux points. Alors que cette relation particulière se révèle importante au niveau linguistique, il devient nécessaire de la définir d'une autre manière, pour qu'elle permette par exemple de tenir compte du temps présent (français) qui peut exprimer des situations dans le passé proche aussi bien que dans des futurs plus ou moins lointains. Finalement, le modèle d'Hornstein est important au niveau conceptuel, surtout parce qu'il intègre le temps d'énonciation et la signification temporelle de l'énoncé en une même représentation.

Un autre travail très important au niveau du traitement linguistique du temps apparaît dans le livre de F. Nef, *Sémantique de la référence temporelle en français moderne* (84) qui traduit une des rares tentatives d'analyse exhaustive du problème pour la langue française. Bien que F. Nef utilise, comme à son habitude, un formalisme logique pour représenter les informations temporelles contenues dans la langue, il s'en sert surtout comme d'un outil méthodologique, ce qui permet d'utiliser ses résultats dans un cadre beaucoup plus général.

En dehors de tout autre aspect, nous discuterons de la structure de temps adoptée par F. Nef. Son modèle repose sur une vision asymétrique du temps entre le passé et le futur, dont les éléments de base sont des événements séparés par des liens assimilables à des relations de causation. Ainsi nous pouvons observer une structure d'événements élémentaires comme dans la figure 2.19.



E0 : Paul tirer au fusil

E : abattre la perdrix.

$\bar{E}$: ne pas abattre la perdrix.

Figure 2.19 : une structure d'événements d'après [Nef 84].

Selon F. Nef, les événements passés se situent sur un axe linéaire du temps du fait qu'ils sont entièrement déterminés, alors que les événements futurs forment une arborescence de choix (Fig. 2.19), qui représente les alternatives existantes à partir d'un événement donné. Ce point de vue est difficilement défendable au niveau cognitif et ceci pour plusieurs raisons :

- le passé, tel qu'il est connu paf un sujet particulier, ou même par un système de dialogue en fonction de son historique, n'est toujours que partiel. Des *trous* subsistent au sein des informations dont nous disposons et ce sont justement ces trous qui sont à l'origine de la Fabula au sens de U. Eco. En effet, nous cherchons constamment à reconstituer l'information qui nous manque en imaginant des situations possibles, qui peuvent s'opposer, partiellement ou totalement. De ce fait, le passé 'cognitif' devient lui aussi une structure où plusieurs choix sont possibles.
- la discussion précédente peut être étendue au futur, dans le cadre duquel nous ne raisonnons pas uniquement en termes d'alternatives à partir de l'instant présent. Lorsque nous désirons atteindre un but, nous étudions un ensemble de chemins possibles pour l'atteindre. La structure alors obtenue ressemble bien plus à un graphe qu'à un simple arbre.

La prise en compte des phénomènes temporels nécessite donc une structure de base plus élaborée que celle proposée par F. Nef, afin de tenir compte de l'activité réelle d'un système intelligent homme ou machine en situation. Dans le cas précis d'un dialogue homme-machine, l'adoption d'une structure de temps linéaire pour le passé, empêcherait le système d'apprendre des éléments nouveaux quant à celui-ci et le conduirait alors à

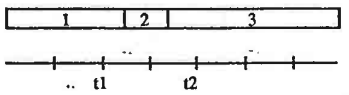
présenter une attitude quelque peu obstinée face à un utilisateur qui en serait décontenancé.

δ Temps et Intelligence Artificielle.

Comme en linguistique, de nombreuses études ont été menées autour des problèmes de représentation du temps en Intelligence Artificielle. Les origines de ces travaux sont de deux ordres :

- la planification, dont les récents développements ont conduit à une prise en compte de la durée et de la synchronisation en environnement réel.
- le traitement automatique du langage naturel, qui est en quelque sorte notre domaine de travail.

Une partie importante de ces travaux, parmi lesquels on peut citer ceux de Mc Dermott (82) et Shoham (87, 88a, 88b), repose sur une représentation du temps dérivée du calcul des propositions ou des prédicats, par adjonction d'instants spécifiant le domaine de validité des formules logiques. Dans ce cadre, le contraste entre un modèle logique et ce que pourrait être un modèle plus cognitif apparaît particulièrement. Si nous prenons par exemple la proposition de Shoham (87) pour une logique temporelle particulière. Sans reprendre ici la syntaxe et la sémantique qu'il introduit (pp.96-98), nous pouvons expliciter à titre d'illustration le prédicat 'TRUE(t1,t2,COLOR(HOUSE17,RED))' qui est une formule bien formé dans cette logique (p.97). Cependant, alors que l'interprétation de cette formule conduit nécessairement à une des deux valeurs vraie ou fausse, la situation réelle présentée dans la figure 2.20 rend cette détermination inacceptable, quelle que soit celle-ci.



intervalle 1 : la maison est rouge.

intervalle 2 : un peintre entre en action.

intervalle 3 : la maison est verte

Figure 2.20 : Un univers indécis pour le prédicat :
TRUE(t1,t2,COLOR(HOUSE17,RED))

Pourtant, répondant à une interrogation de la part d'un inspecteur de police du genre : "Quel était la couleur de la maison entre t1 et t2 ?", tout témoin de bonne foi répondra quelque chose comme : "Ah, mais... la couleur, elle a changé. Un peintre est venu entre temps", etc... ; ce que ne peut pas faire un système reposant sur une logique temporelle, à moins, bien sûr, de transformer à nouveau cette logique pour qu'elle traite cet

exemple...en attendant un nouveau problème. Finalement, ce qui nous intéresse dans un système de dialogue, ce n'est pas tant de connaître la valeur de vérité d'une expression particulière, mais plutôt de posséder des représentations correspondant à une certaine vision du monde et qui soient utilisables avec un maximum de souplesse. De cette manière, l'usager, car c'est toujours lui que l'on cherche à satisfaire, obtiendra un vrai renseignement et non pas un oui ou un non systématique.

Alors que les travaux précédents visaient surtout à ajouter une composante temporelle à des systèmes de déduction déjà existants, Allen (83, 84) a développé une théorie temporelle qui s'appuie explicitement sur des intervalles et donc ne représente que l'information effectivement connue du système. Bien que reposant lui aussi sur une logique temporelle, le modèle d'Allen traite surtout des problèmes spécifiquement temporels. Allen introduit donc une série de treize relations (table 2.2) pouvant lier deux intervalles de manière non-exclusive (plusieurs relations peuvent apparaître entre deux intervalles). Le temps n'est donc défini que de manière relative dans ce modèle, à l'exclusion de toute référence à des bornes numériques. Un intervalle particulier n'est défini qu'à partir des seules relations qu'il établit avec ses voisins, formant de cette manière un réseau de représentation au sein duquel apparaissent certaines contraintes. Afin de gérer ceci, Allen a défini un certain nombre d'algorithmes de propagation de contraintes, qui tiennent compte au passage de nouvelles informations pouvant être connues du système. Ces procédures permettent de plus de vérifier une cohérence temporelle locale au sein d'un ensemble d'intervalles.

Relation	Symbole	Symbole pour l'inverse	Exemple
X before Y	<	>	XXX YYY
X equal Y	=	=	XXX YYY
X meets Y	m	mi	XXXYYY
X overlaps Y	o	oi	XXX YYY
X during Y	d	di	XXX YYYYY
X starts Y	s	si	XXX YYYYY
X finishes Y	f	fi	XXX YYYYY

Table 2.2 : Les 13 (7 + 6 inverses) relations de Allen (83).

Bien que rejetant explicitement la représentation du temps sous la forme de points sur un axe réel, Allen présente ses multiples relations sans s'apercevoir qu'elles introduisent à leur tour les même sortes de problèmes. En effet, quelle est la différence exacte entre les deux relations 'before' et 'meets'? Pour pouvoir répondre à cette question il faut faire intervenir la notion d'intervalle ouvert ou fermé, ou bien, si comme Allen, on cherche à éviter de retomber dans ce travers, développer une sorte de calcul spécifique qui s'appuie sur une notion particulièrement abstraite : le moment ([Allen 85], [Hayes 87]). Alors que ces derniers développements n'apportent strictement rien à la théorie initiale de Allen (bien plus, ils lui font perdre tout son intérêt), il est possible de réenvisager les relations de Allen sous un angle différent qui permettra d'introduire le modèle que nous proposons dans le chapitre suivant.

Le point fondamental que nous voudrions défendre est qu'il est impossible au niveau cognitif de 'penser' à deux événements dont le début ou la fin coïncident exactement au sens mathématique du terme, et même, au sens de Allen. Prenons l'exemple de la précédence. Si nous percevons deux événements bien distincts, nous sommes toujours capable de les situer dans le temps, l'un par rapport à l'autre. Par contre, si ces deux événements se rapprochent, nous ne passons pas brusquement d'un état ou nous les voyons 'before' à un état ou nous les voyons 'meets'. Quand les deux événements sont trop proches pour que l'on puisse prendre une décision, notre attitude est plutôt de raisonner à un niveau plus fin de granularité en observant en détail les extrémités de chaque événement. Nous retrouvons là la notion de profondeur variable chère à D. Kayser.

Une discussion similaire est envisageable pour les relations 'equals' (on considère alors deux bornes) 'starts' et 'finishes'. En définitive, nous pouvons réunir les treize relations de Allen en trois catégories distinctes :

- les relations d'adjacence, parmi lesquelles nous trouvons 'before', 'meets' et leurs inverses respectifs.
- les relations d'inclusions, qui contiennent 'starts', 'during', 'finishes' et leurs inverses.
- la superposition, qui ne contient que 'overlaps' et son inverse.

Alors que les deux premières classes se conçoivent facilement en fonctions de critères cognitifs ([Piaget 46]), la dernière pose plus de problèmes. En fait, contrairement aux autres classes que nous avons introduites, 'overlaps' ne possède pas cette transitivité d'ensemble qui la rendrait facile à percevoir. De plus une discussion similaire à celle qui a

été faite pour grouper les autres relations permet de montrer qu'elle ne se justifie pas comme relation fondamentale d'un modèle de représentation d'informations temporelles. Ainsi, il semble qu'à nouveau, la prise de conscience de deux événements qui se superposent passe d'abord par une étude, à un niveau inférieur, de la zone commune entre ces deux événements, ce qui entraîne l'utilisation des relations de précédence et d'inclusion pour exprimer une telle situation.

2.7.2 Conclusion.

De l'étude des différents modèles de représentation du temps, nous pouvons tirer quelques points intéressants pour notre recherche. En premier lieu, l'étude, même partielle des logiques temporelles, ne nous a pas réconcilié avec les approches formelles que nous avions déjà critiquées. Dans un deuxième temps, nous avons trouvé chez Allen de très intéressantes indications pour un réel modèle cognitif du temps, bien que certaines de ses relations semblent redondantes. L'originalité de ce modèle qui se situe effectivement à l'opposé de toute représentation numérique explique tout à fait la popularité qu'il a conservée depuis et les nombreux travaux qui se fondent sur ces relations ([Lâasri 88], [Granier 88], [Ladkin 87], [Kumar 87], [Tsang 87]).

3 UN MODELE COGNITIF DE REPRÉSENTATION DU TEMPS.

Nous proposons dans cette partie un modèle de représentation des informations temporelles dans un dialogue oral homme-machine, pouvant par extension traiter différents problèmes liés au langage en général. Après avoir précisé nos objectifs et les principaux éléments de notre modèle, nous détaillerons un exemple complexe qui explicitera les principales possibilités de représentation dont nous disposons. Ensuite, nous introduirons les outils supplémentaires par lesquels le modèle peut opérer. A partir de cela, nous verrons comment ce modèle s'applique dans différents domaines. Enfin nous essayerons de faire un bilan des bénéfices que nous pouvons tirer de cette approche, ainsi que de la réflexion qui reste à mener.

3.1 OBJECTIFS DE CE MODELE.

3.1.1 Un modèle pour représenter.

Nous avons vu au chapitre 2.7 que le temps, dans tout échange linguistique, prenait des formes très diverses. Parmi celles-ci un système doit prendre en compte plus particulièrement :

- le temps de l'énoncé. Il s'agit d'un temps perceptuel qui va éventuellement donner au système des points de repère autour desquels il pourra situer les autres objets de son univers.
- le temps des événements du discours. Chaque énoncé décrit une situation se déroulant dans le temps et qu'il faut à ce titre représenter. Réaliser cette opération pose de manière plus générale le problème de l'association d'un énoncé à son contenu sémantique.
- le temps du raisonnement, qui apparaît dès que le système manipule les objets de l'univers de sa tâche qui doivent garder une cohérence temporelle tout au long d'une session de dialogue par exemple. En particulier, le système doit pouvoir représenter les différents stades de connaissance d'un usager en quête d'informations.

Ces différentes façons de considérer le temps ne sont pas incompatibles avec une représentation commune. Au contraire, un bon modèle de représentation du temps ne sera cohérent que s'il possède une structure homogène pour toutes ces informations.

3.1.2 Un modèle pour apprendre.

Notre but n'est pas seulement de représenter des informations temporelles brutes à partir d'un dialogue, mais il est important de structurer ces informations pour dégager des schémas abstraits qui seront réutilisables. Ce n'est qu'à cette condition que nous obtiendrons cette souplesse de fonctionnement que nous avons exigée des théories étudiées au chapitre 2.

L'apprentissage que nous désirons voir fonctionner dans un système de dialogue homme-machine, doit tout d'abord assurer la formation d'un contexte du discours. Ce contexte concerne tout aussi bien les données linguistiques que les données pragmatiques, c'est à dire les objets dont chaque énoncé a parlé. La mémorisation de tous ces éléments doit éviter les pièges d'une structuration par niveaux, du syntagme au dialogue, car cela entraînerait, comme nous l'avons vu au 1.4, une difficulté pour combiner ces informations.

Le système doit aussi prendre du recul par rapport aux informations qu'il a acquises, pour proposer des motifs temporels qui se retrouvent dans le discours. Cet apprentissage ne doit pas uniquement concerner les objets du discours, mais aussi les structures de la langue qui sont très variables dans un dialogue oral. La compréhension d'un énoncé 'inhabituel' en sera ainsi facilitée.

3.1.3 Un modèle pour prédire.

Une caractérisation possible d'un système intelligent peut résider dans son aptitude à anticiper sur les évolutions de son environnement. Ceci doit lui permettre d'une part de compléter les informations qui lui manquent, d'autre part, de sélectionner les informations applicables en fonction d'un contexte de raisonnement donné, afin d'optimiser l'acquisition de nouvelles connaissances. Dans le cadre d'un dialogue oral homme-machine ce processus de prédiction concerne deux domaines principaux.

Au niveau linguistique, une prédiction consiste à générer automatiquement des éléments de la langue susceptibles d'apparaître dans un énoncé, soit dans le cas où ces éléments sont fortement bruités et donc méconnaissables, soit simplement par anticipation ; ce qui permet au système de n'avoir qu'à vérifier ces éléments plutôt que de les rechercher. Ce phénomène est particulièrement apparent dans le cas de la lecture de

textes ([Dubois 88]) où il a été prouvé que certains mots ne sont même pas regardés quand ils n'apportent aucune information nouvelle au lecteur.

Un deuxième mode de prédiction concerne les événements en général dans l'univers de raisonnement. En effet, un système doit prévoir différentes situations dans le domaine de la tâche pour éviter, par exemple, de conseiller à un usager de se rendre à la mairie si celle-ci doit fermer d'ici peu et lui préconiser alors le commissariat de police. Ce phénomène est d'ailleurs beaucoup plus général que cela, si on désire qu'un système de dialogue comprenne effectivement chaque énoncé du locuteur. Comme dans le cas de la compréhension d'un texte écrit ([Eco 85]), la compréhension d'un dialogue implique une activité de fabulation qui positionne chaque énoncé dans une structure discursive cohérente. Il faut ainsi prévoir que si l'usager déclare avoir perdu sa carte d'identité, il va sûrement demander des renseignements concernant les modalités de son remplacement, sachant que ce papier est quasiment indispensable. C'est alors que l'on pourra parler effectivement de comportement intelligent.

Enfin, comme nous l'avions observé pour les facultés de représentation, les modes de prédiction précédents sont assimilables à une seule et même opération dans le cadre d'un modèle du temps. Ainsi, le système sera à même de combiner des prédictions syntaxiques et sémantiques pour arriver à un résultat unique.

3.1.4 Un modèle cognitif.

Quels sont les critères qui permettent de qualifier un modèle de représentation de 'cognitif' ? Je pense qu'ils sont assez difficiles à établir. Néanmoins, nous pouvons donner quelques indications sur ce que nous considérons comme important pour qu'une telle propriété soit vérifiée.

Quand Yip (85) parle de représentation cognitive du temps, il présente une manière de traiter des problèmes que l'homme manipule avec facilité. Plutôt que de faire une liste de ces problèmes, il montre qu'avec un certain nombre d'outils de base, on peut directement les cerner. En conséquence, il considère qu'il est beaucoup plus important d'expliquer les phénomènes plutôt que de les décrire, ce en quoi nous sommes d'accord, si ce n'est qu'il est nécessaire de décrire correctement un objet avant de songer à le comprendre.

Un deuxième aspect, que nous avons déjà évoqué (2.6), est qu'un modèle cognitif satisfaisant s'oppose d'une certaine manière à un modèle logique et plus généralement

trop formel. Il y a là matière à débat, pour notre part, l'approche que nous avons choisie nous conduit à rechercher un modèle cognitif en dehors des modèles logiques classiques.

Finalement, le point le plus important pour nous est de reconsidérer la situation qu'occupe le système dans le couple homme-machine. Les derniers développements de la recherche cognitive nous montrent que l'on s'éloigne du schéma de la théorie de la communication pour analyser ce qu'est un locuteur-auditeur en situation. Ainsi, plutôt que d'observer un système de dialogue de l'extérieur, en essayant de programmer un comportement qui au bout du compte soit satisfaisant, il devient important de réellement analyser ce que pourrait être son fonctionnement interne (ou 'cognitif') de façon plus abstraite. Ce n'est pas en ajoutant des procédures que l'on améliorera vraiment le naturel du comportement d'un système. Finalement, faire du dialogue homme-machine passe par la volonté de réaliser un système intelligent.

3.2 Les premiers éléments.

3.2.1 Les zones temporelles.

1 Définition.

Toute entité susceptible d'être manipulée par un système intelligent et se déroulant dans le temps est représentée dans ce modèle par une *zone temporelle*. Les zones temporelles sont les seuls objets manipulés explicitement par ce modèle.

2 Commentaires.

α Expressivité.

Nous n'avons introduit qu'un seul objet pour représenter de nombreuses entités qui sont souvent différenciées. Le fait de considérer une unique catégorie permet d'envisager un traitement unifié de tous ces objets.

Les objets apparaissant comme zones temporelles peuvent provenir de différentes origines, si on considère un système intelligent de manière générale. La principale, et je dirais initiale, source de formation de zones temporelles est la perception. Ainsi, le fait de voir passer quelque chose devant son champ visuel entraîne la création d'une zone associée à cette perception, qui traduit ainsi son homogénéité dans le temps. Une zone est

donc la prise de conscience par un sujet qu'un événement forme un tout qui peut être assimilé à un seul concept, à un niveau donné. De ce fait, les zones temporelles ne distinguent pas, a priori, les événements dynamiques par opposition à des situations figées. Percevoir ces dernières correspond en quelque sorte à reconnaître une identité de perception au cours du temps. Il n'y a donc qu'une différence quantitative entre une situation qui évolue et une situation où l'évolution est proche de zéro, d'où notre choix de ne pas les distinguer. Inversement, un événement particulier, comme la vision d'une étoile filante peut être vue comme un tout, à un niveau donné d'abstraction, sans affiner plus la structure temporelle. Là encore, la représentation se fera par une seule zone temporelle.

Parmi les perceptions les plus courantes, une certaine catégorie d'entre elles nous intéresse plus particulièrement dans un dialogue oral homme machine, ce sont les perceptions auditives, que l'on qualifiera ici plutôt de langagières. Ainsi, un phonème, un mot ou un énoncé tout entier peuvent être considérés comme des ensembles cohérents d'informations se déroulant dans le temps et donc représentés comme des zones temporelles dans notre modèle. Nous considérerons donc un énoncé, ou une quelconque partie de celui-ci, comme des événements à part entière.

En marge des zones perçues, les zones temporelles qui peuvent être générées dans le cadre de ce modèle sont les zones signifiées (ou racontées). La présence d'une certaine zone dans l'espace de représentation induit, par association, la création d'une autre zone qui lui est associée. Ainsi toute phrase simple entraîne la création d'une zone correspondant à l'action (situation, événement, processus ou autre) décrite par cette phrase. Soit par exemple l'énoncé : "je chante". Nous venons de voir qu'il est représenté dans notre modèle par une zone temporelle que nous schématiserons ainsi :

(les guillemets indiquent qu'il s'agit d'un message perçu)

Or cette énoncé traduit une situation qui possède un certain développement dans le temps que nous représenterons sommairement ainsi :

Il apparaît donc une association entre certaines parties d'un énoncé et des représentations abstraites qui en sont faites. Nous les représenterons de la façon suivante avant de détailler un peu plus ce mécanisme (3.4) :

Ce qu'il est important de constater ici c'est que cette association peut être réalisée à de nombreux niveaux d'une phrase, du mot ("hier") jusqu'à un énoncé tout entier, en passant par des locutions diverses ("la nuit de samedi à dimanche"). A ce type d'expressions est associée une seule zone si elles peuvent être perçues comme formant un seul concept. Par exemple l'expression "les trois derniers jours" décrit un intervalle de temps qui peut être vu comme un tout (une seule zone). Eventuellement, une phase de raisonnement découpera cette zone en trois sous-zones plus spécifiques.

Nous éviterons, dans la mesure du possible, d'étendre cette représentation aux objets "réels" qui sont habituellement manipulés par les systèmes de représentation des connaissances. Une table par exemple, comme objet perçu, possède un déroulement temporel et l'on peut facilement envisager qu'au message "la table" on associe une zone qui représente l'existence de celle-ci à travers le temps. Nous nous limiterons d'abord aux messages *purement* temporels tels que "dimanche", "une journée", "je marche", etc...

Un dernier point sur lequel nous désirons insister est qu'une zone temporelle n'est que la projection de ce que pourrait être un concept *plus complexe* dans un système de représentation général. Tous les problèmes traités dans ce chapitre le seront dans l'optique du temps à l'exclusion de toute autre considération qui pourrait concerner une extension possible de ce modèle.

β Valeur temporelle.

Lorsque nous pensons à un événement dont nous avons le souvenir, il est souvent difficile de définir avec certitude son commencement et sa fin. Bien plus, si nous cherchons effectivement un événement dont nous pourrions donner les bornes précises, nous sommes quasiment incapable d'en trouver un. La perception que nous avons d'une situation ne peut donc être calquée sur un quelconque axe numérique qui nous servirait de référence. Partant de ce point de vue, les zones temporelles sont des intervalles de temps sans aucune borne ce qui les rapproche de la vision du temps qu'a Allen.

Soit par exemple l'expression : "la nuit dernière". Alors que tout individu est capable de raisonner convenablement sur le concept associé à cette expression et de produire des

énoncés tels que : "la nuit dernière, il y a eu un gros orage", il sera ennuyé si on lui demande de préciser pour quels instants particuliers il considère que l'expression "la nuit dernière" a été valide. Tout au plus, il pourra indiquer des situations extrêmes où il est sûr que *c'était la nuit dernière*, ou que ce ne l'était pas. Si par hasard l'orage s'est produit un peu plus tôt de sorte que "la nuit dernière" ne permette plus de le situer, il usera d'une autre expression plus appropriée telle que "hier, dans la soirée". Nous pouvons représenter pour l'instant cette situation par la figure 3.1 en comparant toutes ces zones avec un éventuel axe temporel.

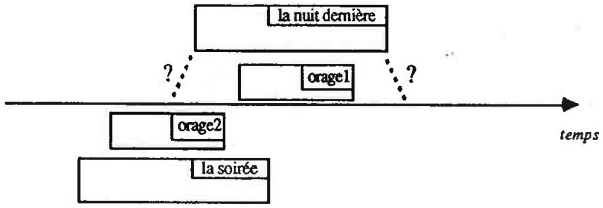


Figure 3.1 : Zones temporelles imprécises versus axe temporel précis.

Une objection peut être faite aux affirmations avancées ci-dessus : il existe malgré tout des montres et des calendriers dans notre univers de tous les jours et l'expression "trois heures trente" a finalement un sens. Effectivement, nous disposons d'outils de mesure du temps qui peuvent être très précis. Mais cette précision a-t-elle un sens au niveau cognitif ? Je ne le pense pas. Si nous consultons des horloges, c'est justement parce que nous n'avons pas de moyen de déterminer le temps de manière absolue. A chaque regard sur notre montre, nous mémorisons un point de repère qui va servir à positionner les autres événements que nous percevons. Cette "faiblesse" qui nous caractérise présente quand même un avantage qui peut être réutilisé dans le cadre de la mise en œuvre de notre modèle sur une machine. En effet, il apparaît un phénomène d'économie de représentation vis à vis du temps quand on ne mémorise pas continuellement toute une série de données numériques liées à un événement particulier.

Pour reprendre l'exemple de l'expression "trois heures trente", nous n'associons pas à celle-ci un laps de temps constant qui serait à la limite ponctuel. Contrairement à ce qu'affirme A. Borillo (86 : note 1 p.15), nous associons à cette expression une zone comme les autres qui dure obligatoirement le temps de sa perception, dans le cas où l'on dit "il est trois heures trente". Dans un autre contexte, si l'on donne un rendez-vous à quelqu'un à "trois heures trente", on pourra considérer que la personne en question est

ponctuelle si elle arrive dans une période correspondant à un ensemble de lectures d'horloge autour de l'indication 3h30. Cette période peut être très variable si l'on n'a pas d'aiguille des secondes pour soutenir la pensée. Enfin, un technicien d'Ariane Espace fera correspondre dans son univers cognitif l'expression "trois heures trente" au temps où son horloge atomique indiquera cette mesure. La durée est donc, tout comme la position dans le temps, une donnée extrêmement relative.

Dans le cas d'un dialogue homme-machine, le système possédera assez peu de points de repères. En fait, il n'aura qu'une seule source de renseignements au niveau temporel à travers les différents énoncés qui lui parviendront de la part de l'usager. Le positionnement des différents objets du discours se fera alors suivant le même type de schéma d'analyse que celui proposé par Hornstein (77). Donner au système une horloge interne très précise n'apportera finalement rien à la qualité des échanges, au contraire, le locuteur n'a que faire d'une machine qui ne raisonne pas sur les mêmes objets que lui.

γ Différenciation des zones.

Puisque toutes les entités manipulées par notre modèle sont représentées par un unique objet, il faut lui donner des moyens de différencier chaque zone particulière par rapport aux autres. En dehors des problèmes de perception et de génération sur lesquels nous reviendrons ultérieurement, les zones temporelles ne sont distinguées que par les relations qu'elles établissent entre elles. Conformément à notre vision d'un temps uniquement relatif, il faut en effet que nous définissions des liens temporels entre zones, qui permettent de les situer les unes par rapport aux autres. De plus nous verrons qu'il est nécessaire d'introduire un mécanisme d'héritage afin de mémoriser des structures temporelles standard dans l'univers de représentation. Dans un premier temps nous allons détailler les relations temporelles utilisées par notre modèle.

3.2.2 Les relations entre les zones.

1 INTRODUCTION.

Au moment de la présentation du modèle de Allen (2.7.2), nous avions mis en avant les redondances que présentaient ses relations temporelles et nous avions montré que celles-ci pouvaient être réduites à trois classes indépendantes, à savoir les relations d'adjacence, d'inclusion et de superposition. A l'intérieur de ces classes et plus particulièrement pour les deux premières, la différence qui était faite entre des relations

orientées dans le même sens ne pouvait pas être justifiée dans une optique cognitive. Parallèlement, la dernière relation - la superposition (*'overlaps'*) - n'exprime pas des phénomènes qui sont couramment rencontrés tels quels dans le raisonnement de tous les jours. Aussi nous n'introduisons que *deux relations* temporelles entres les zones que nous avons introduites : l'*adjacence* et l'*inclusion*.

Indépendamment de la comparaison avec le modèle de Allen, ces deux relations se justifient comme étant celles qui sont manipulées le plus naturellement au niveau cognitif ([Piaget 46]). De plus leur faible nombre et donc leur simplicité, permet d'envisager des traitements réduits tant au niveau de la représentation, que des vérifications de cohérence parmi un ensemble de zones qui seraient connectées de la sorte.

2 L'ADJACENCE.

α Sémantique.

Deux zones Z1 et Z2 sont adjacentes (noté : *prec(Z1,Z2)* ou *prec_i(Z2,Z1)*) si Z1 est perçue par un système de représentation avant Z2 de manière distincte au niveau temporel.

Schématiquement cette relation sera représentée ainsi : 

β Compléments.

Cette relation permet de représenter toute suite d'événements et donc de situer explicitement un événement par rapport à un point de repère dans le passé ou le futur. Dans un dialogue par exemple, les différents échanges entre la machine et le locuteur ne formeront pas un ensemble d'informations en vrac, mais seront liés entre eux par des liens d'adjacence qui les localiseront dans le temps.

Si par exemple, l'usager demande des renseignements sur les formalités à remplir pour un mariage, le système est en mesure de raisonner par rapport à l'événement "mariage" et produire des renseignements sur la base de celui-ci : "Avant votre mariage, vous devez...".

Cette relation n'a pas seulement une valeur temporelle si on la considère de façon un peu plus abstraite. En effet, nous avons vu avec Piaget que l'expression de la cause

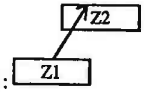
pouvait résulter de la prise de conscience d'une reproductibilité de certains schémas temporels. Cette idée est reprise notamment par Lyons (77 : p.493). Plus généralement l'observation, par un système, de suites d'événements qui se retrouvent régulièrement va entraîner la mémorisation de celles-ci sous la forme de zones standards liées par des relations d'adjacence. En plus de la causation, on peut songer à représenter de la sorte différents phénomènes du type symptôme-maladie, observation-prédiction, etc, dès que ces ensembles de situations se succèdent dans le temps.

3 L'INCLUSION.

α Sémantique.

Une zone Z1 est incluse dans une zone Z2 (noté : *in(Z1,Z2)* ou *in_i(Z2,Z1)*) si Z1 est perçue par un système de représentation à l'intérieur de Z2 de façon distincte.

Schématiquement cette relation sera représentée ainsi :



β Compléments.

De même que pour la relation d'adjacence, on peut envisager sous deux aspects différents cette relation d'inclusion :

- au niveau purement temporel, elle permet de représenter des situations explicitement imbriquées, comme dans la phrase "Je chantais quand Léon est entré" de la manière schématisée figure 3.2, où nous avons au passage situé l'événement par rapport à l'énoncé. Afin de différencier les zones de "texte" des zones construites, nous avons étiqueté ces dernières par des formes prédicatives qui les représentent sans que ce point ait une grande importance pour l'instant.

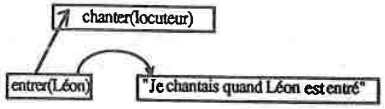


Figure 3.2 : une représentation de l'énoncé :
"Je chantais quand Léon est entré".

Un deuxième aspect que peut recouvrir la relation d'inclusion concerne la description qui peut être faite d'un événement particulier suivant différents niveaux de précision. Ainsi, si une zone Z1 est incluse dans une zone Z2, ceci peut indiquer que Z1 est une composante de la zone plus générale Z2. Prenons un exemple. Lorsque parlant de son activité de la veille, une personne asserte "hier, j'ai marché toute la journée", elle exprime un état de fait qui, tel qu'il est décrit dans cet énoncé, peut être représenté par une zone temporelle correspondant grossièrement à une journée. Cependant, la personne n'a peut-être pas effectivement *marché* toute la journée au sens propre du terme. Si on lui demande de préciser ce qu'elle a fait, elle pourra dire notamment : "je me suis arrêté pour manger". Cette situation peut être schématisée par la figure 3.3 où les portions d'énoncés référants des zones précises sont rapportées.

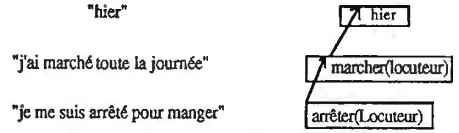


Figure 3.3 : représentation d'une promenade.

Chaque zone qui est représentée dans cette figure ne doit pas être prise au sens littéral de l'énoncé correspondant, ce qui conduirait bien-sûr à une incohérence par rapport à deux actions différentes que ferait le locuteur pendant une même période. Ce qu'il est important de comprendre, c'est que ces zones sont une matérialisation d'un ensemble d'informations temporelles cohérentes, à un certain niveau. De manière plus précise, la nécessité d'exprimer, par le langage, certains de ces groupements, focalise l'attention du locuteur sur ceux-ci. Inversement, le système qui entend les expressions linguistiques correspondantes, construit les zones associées puisque le locuteur les déclare pertinentes. Cependant, alors que le locuteur *dispose* des informations de détail relatives à ces zones, l'auditeur ne peut travailler que sur ce qu'il a entendu. Si l'auditeur désire préciser ces zones, il devra par exemple mettre en œuvre des schémas qu'il possède relatifs à une marche *typique* où il est nécessaire de s'arrêter pour manger.

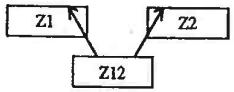
Nous apercevons ici une des propriétés (défaut ?) essentielle de la langue qui est de ne donner qu'une description sommaire d'une réalité qui, pour le locuteur en particulier, peut s'avérer beaucoup plus complexe. La langue n'exprime pas tout et parfois elle n'exprime même pas une proposition que le locuteur prendrait pour vraie comme par exemple (en logique temporelle approximative) : {∀t ∈ [Hier], marcher(Locuteur,t)}. On retrouve alors la nécessité pour l'auditeur (le système) d'avoir une activité de Fabula, pour ne pas

se voir obligé de raisonner sur des informations trop grossières et déduire une incohérence dans la situation exposée ci-dessus.

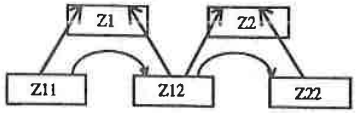
4 PROPRIÉTÉS.

Une première propriété de ces deux relations est qu'elles sont mutuellement exclusives. Cela signifie qu'entre deux zones données Z1 et Z2 on ne peut trouver qu'une de ces deux relations ou pas de relation quand aucune information n'est disponible. Ce point est important, car il n'y a ainsi jamais d'ambiguïté de représentation entre deux zones comme cela existe dans le modèle d'Allen où il maintient un ensemble de contraintes correspondant aux relations encore possibles entre deux intervalles. Ici les seules contraintes se résument à déterminer *la* relation qui existe entre deux zones. Nous verrons par la suite que cela accélère aussi le processus de vérification de la cohérence temporelle au sein d'un ensemble de zones.

Nous pouvons maintenant revenir sur la relation de superposition que nous avions rejetée car elle ne nous semblait pas fondamentale au niveau d'un système de représentation d'informations temporelles. La superposition entre deux zones traduit en fait une situation où Z1 et Z2 ont en commun une troisième zone Z12 (au sens de l'inclusion) Nous pouvons représenter cette situation ainsi :



Dans ce cas, la représentation n'est pas strictement équivalente à la superposition au sens de Allen, car les zones Z1 et Z2 peuvent être incluses l'une dans l'autre sans qu'il y ait une incohérence temporelle. Si on désire préciser que Z1 et Z2 ont chacune des parties qui leur sont propres, il faut décrire plus finement leurs extrémités respectives en introduisant par exemple des zones Z11 et Z22 de la façon suivante :



Il est ainsi possible d'obtenir avec les éléments de base de notre modèle une représentation satisfaisante du '*overlaps*' de Allen qui est de plus tout à fait

compréhensible quant à son origine possible, quand on cherche à analyser deux situations, l'une par rapport à l'autre. Notre modèle a donc une puissance de représentation au moins égale à celle de Allen dans le cadre d'un système de raisonnement de sens commun, tout en utilisant un nombre bien inférieur de relations.

3.2.3 Introduction à l'étude d'un exemple.

Avant d'aborder l'étude d'un exemple destiné à se familiariser avec les zones temporelles, nous introduisons deux notions intermédiaires qui serviront dans l'analyse à venir :

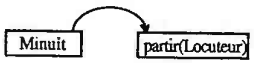
- les *zones d'attention*, que nous nommerons parfois *zones d'ancrage.*
- quelques indications sur les marqueurs linguistiques utilisés pour générer des schémas temporels.

1 LES ZONES D'ATTENTION.

Soit l'expression "Après minuit,..." ; nous pouvons facilement lui associer une zone temporelle pour la sous-expression "minuit", mais il est plus difficile de définir une structure précise autour de cette zone. En fait, la seule chose dont on peut être sûr, c'est l'existence d'une autre zone qui lui succède sans qu'on possède plus de détail sur la nature exacte de celle-ci (de quel événement il s'agit par exemple). De manière générale, il arrive souvent que nous puissions faire une prédiction relative à l'apparition d'une zone temporelle, sans pouvoir en déterminer l'origine. C'est pour représenter ce phénomène que nous avons introduit le concept de zone d'attention (ou d'ancrage) qui correspond à une attente d'information complémentaire pour une zone qui a été prédite mais qui reste néanmoins inconnue. Ainsi nous représenterons l'expression "Après minuit,..." comme suit :

Les zones d'attention correspondent donc à des sortes de prédictions floues prêtes à être remplacées par la première information disponible. Elles vont donc servir en particulier à relier différentes expressions, propositions ou phrases, en permettant de compléter grâce à l'une d'entre elles, l'information qui était attendue dans une autre. Si en

complément de "Après minuit", on a ainsi "je suis parti", l'action décrite par cette dernière proposition trouvera naturellement sa place :



Bien sûr, l'information complémentaire peut se situer avant ou après l'apparition de la zone d'attention, comme dans "Je suis parti après minuit". Ce mécanisme est relativement général et permet de retrouver la continuité d'un discours ou même d'un dialogue, puisque ce peut être un moyen de comprendre le lien entre la question : "Quand es-tu parti ?" et la réponse : "hier". Nous n'avons pas encore développé ces voies de recherche liées aux actes de langage, car elles dépassent pour l'instant le cadre de notre analyse.

2 Les marqueurs temporels.

Dans l'exemple qui va suivre, nous allons construire des représentations temporelles à partir de seuls marqueurs linguistiques présents dans le texte à analyser. En attendant d'étudier dans le détail le mécanisme concerné, le principe que nous adopterons consistera à associer directement à certaines formes précises, des schémas temporels que nous essayerons de lier les uns aux autres ; ce que nous représenterons parfois sous la forme :

objet linguistique → *schéma temporel*

Les principales classes de marqueurs linguistiques à valeur temporelle que nous utiliserons seront les suivantes :

- les prépositions : *dans, après, à ...*
- les conjonctions de subordination qui relient explicitement deux propositions : *quand, après, pendant que...*
- les désinences de temps de conjugaison, qui situent dans le temps les événements décrits.
- certains items lexicaux à valeur temporelle importantes : *hier, dimanche...*

Leurs différentes valeurs seront précisées au fur et à mesure de leurs apparitions.

3.3 Analyse d'un exemple : Le nom de la rose.

3.3.1 Introduction.

Extrait du roman : *Le nom de la rose* de Umberto Eco (Grasset).

> Premier jour
> PRIME
> Où l'on arrive au pied de l'abbaye
> et Guillaume fournit une preuve
> de sa grande sagacité.
>
> C'était une belle matinée de la fin novembre. Dans la nuit, il avait neigé un peu, mais le terrain était recouvert d'un voile frais pas plus haut que trois doigts. En pleine obscurité, sitôt après laudes, nous avions écouté la messe dans un village de la vallée. Puis nous nous étions mis en route vers les montagnes, au lever du soleil.
>
> Comme nous grimpions par le sentier abrupt qui serpentait autour du mont, je vis l'abbaye.

Il s'agit d'un texte narratif présentant une partie des temps usuels du passé (Imparfait, Passé simple, Plus-que-parfait) et dont l'analyse fait apparaître des éléments importants pour tester le modèle proposé. Il présente la particularité, que l'on retrouve dans de nombreux récits, de situer une scène complexe avant d'introduire l'action principale, qui marque le début de l'histoire effective : "je vis l'abbaye". Le principe de l'interprétation repose à tout moment sur l'existence d'une pseudo-zone temporelle correspondant au point courant d'attention. Tout nouvel événement va prendre la place de cette zone ou se situer par rapport à celle-ci, et, en fonction de ses caractéristiques temporelles propres, (temps verbal de l'énoncé, connecteurs temporels etc...) créer un nouveau point d'ancrage pour les énoncés futurs. Localement, l'interprétation d'une phrase particulière peut nécessiter l'introduction d'une zone temporaire qui devra être située par rapport à la zone précédemment active. Les marques linguistiques intervenant en cours d'analyse seront citées au fur et à mesure, ainsi que les raisonnements qu'il serait nécessaire de mettre en œuvre pour effectivement *comprendre* l'énoncé. Nous ferons entre autres apparaître des phénomènes d'intertextualité (utilisation d'anciennes connaissances extraites de lectures précédentes) qui sont fondamentales pour la compréhension au sens de Eco (65, 85).

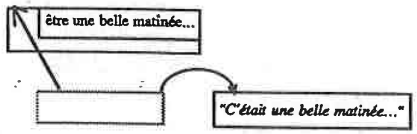
3.3.2 Analyse linéaire et construction d'une représentation.

Un texte écrit diffère profondément de l'oral en ce sens que la réception du message est retardée par rapport à son émission. Une bonne attitude d'interprétation est de considérer que le temps de l'énoncé est le moment où le texte a été produit. Le problème n'est pas crucial ici, du fait qu'il nous suffit de posséder un point de repère par rapport auquel situer l'histoire dans le passé.

On commence donc la lecture avec en mémoire une zone d'ancrage, dont la position est indéterminée dans le temps, au sens où l'action peut se situer dans le passé, le présent ou l'avenir, indifféremment :

"C'était une belle matinée de la fin novembre."

Indépendamment de la sémantique de "c'était" et de "une belle matinée...", On peut schématiser la formule ci-dessus comme étant une proposition "ce être une belle matinée..." conjuguée à l'imparfait ce qui situe le point d'ancrage courant à l'intérieur de l'action correspondante. L'énoncé instancie donc une nouvelle zone temporelle (le temps de la situation décrite par la phrase) dans laquelle est incluse la zone de référence. On obtient alors une structure relativement simple schématisée ainsi :



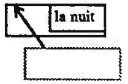
Il est dès maintenant utile d'émettre certaines remarques relatives à la structure ci-dessus. Tout d'abord, il n'est pas possible de connaître les liens temporels entre la zone "être une belle matinée..." et celle de l'énoncé. Localement, on peut mettre en route un système de raisonnement qui essaie d'établir une des relations possibles entre ces deux éléments, en s'appuyant alors sur des connaissances plus précises de l'univers (en particulier, le prologue du roman indique que le temps d'énonciation se situe à la fin de la

vie du narrateur, et que l'histoire, elle, se déroule alors que celui-ci était adolescent). Cependant la situation où la zone désignée contient à la fois le temps de l'énoncé et le temps de narration est aussi possible, comme dans "C'était cette semaine...". Pour aller plus loin dans cette analyse, il sera nécessaire de détailler la représentation associée à la sémantique détaillée de la phrase.

"Dans la nuit,..."

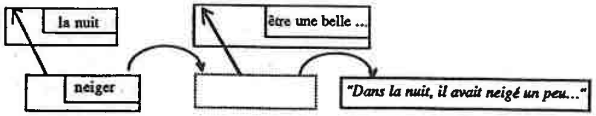
Nous arrivons sur une structure syntaxique particulière qui déclenche la création d'un point d'ancrage temporaire à l'intérieur de la zone induite par le corps de cette structure. On pourra représenter ceci au niveau lexical, en associant à l'entrée "dans" un fragment de représentation, destiné à être activé pour une certaine construction dans laquelle ce mot interviendrait. Ce qui peut être schématisé ainsi :

On obtient donc la structure locale suivante :



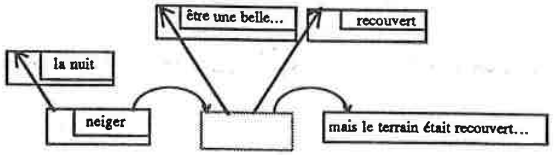
"...il avait neigé un peu,..."

Cette proposition s'insère immédiatement dans le point d'ancrage qui vient d'être généré, et se trouve située par rapport au point d'attention laissé en suspend précédemment, par l'intermédiaire du plus-que-parfait qui indique une action précédant le processus en cours de description. Localement, il faut que les zones temporelles désignées par "Dans la nuit" et le plus-que-parfait soient cohérentes, ce qui nécessite quelques inférences avant de procéder à l'unification des structures. Ainsi la phrase "Demain, il pleuvait" ne peut être interprétée car elle possède des schémas temporels contradictoires. D'où une représentation partielle :



"...mais le terrain était recouvert d'un voile frais pas plus haut que trois doigts."

Une nouvelle proposition est représentée ("recouvert(terrain, neige)") qui, du fait de l'imparfait, contient le point courant (on se situe à l'intérieur de l'action, du processus, ou même d'un état qui dure). On ne peut positionner cet état par rapport à "matinée". De fait, l'observation de la présence de neige n'est dans un premier temps importante que très localement, mais il est possible que celle-ci soit utilisée bien plus tard (au delà de la matinée) dans un raisonnement. Cependant, cette représentation permet de faire ressortir le fait que les deux zones "matinée" et "recouvert" se superposent, au moins partiellement, au niveau de la zone active commune à chacune d'elles. Enfin, on ne précise pas ici la valeur de "mais", car il n'intervient pas directement au niveau temporel. Il traduit plutôt une opposition vis à vis d'un processus d'interprétation normal de l'énoncé, notion qui reste à développer. La structure obtenue jusqu'à maintenant donne :

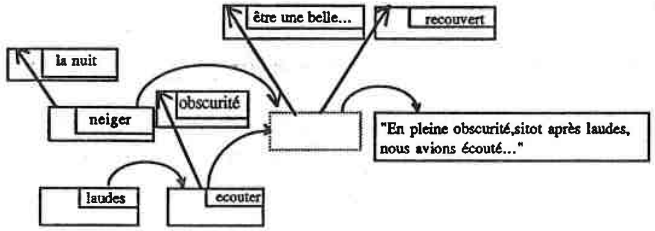


On remarque encore certaines déductions qu'il serait localement bon de faire, comme par exemple établir le lien entre la neige sur le sol (citée de façon métaphorique ici), et le fait qu'il ait neigé. Ceci peut s'appuyer sur des schémas temporels typiques, relatifs à certaines actions et à leurs conséquences.

"En pleine obscurité, sitôt après laudes, nous avions écouté la messe dans un village de la vallée."

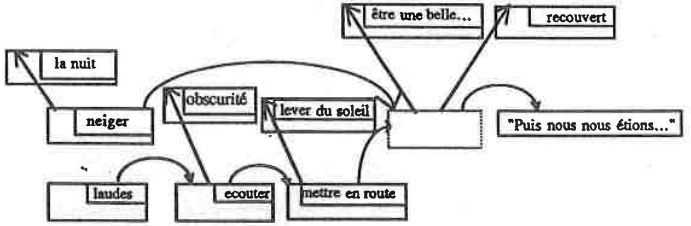
On retrouve la plupart des éléments déjà cités : présence du plus-que-parfait situant une action antérieure au point courant, et établissement d'une relation locale par rapport à

"laudes" (une des heures bénédictines) et par rapport à "la pleine obscurité". Cette dernière zone est un peu particulière. En effet, il est difficile de définir exactement son caractère temporel ou spatial. Cependant, la relation d'inclusion de l'objet "la pleine obscurité" par rapport à l'action courante peut être maintenue moyennant un certain flou sur le type exact de zone dont il s'agit. On néglige dans un premier temps l'aspect contenu dans "sitôt", qui devra être interprété ultérieurement dans une version plus complète du modèle. On donne donc la représentation correspondant à la phrase :



"Puis, nous nous étions mis en route vers les montagnes, au lever du soleil."

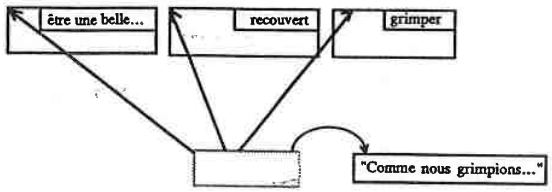
"Puis" situe l'action courante par rapport à la dernière rencontrée (qui peut être retrouvée par la succession des énoncés toujours disponible). Associée au plus-que-parfait et positionnée par rapport "au lever du soleil", cela permet d'établir la structure suivante :



Comme nous l'avons remarqué, après l'interprétation de chaque énoncé, il est nécessaire de lancer une phase de raisonnement sur les éléments créés, afin de déterminer certaines relations non explicitées dans le texte, mais déductibles de connaissances générales du monde. (Ainsi il est dit dans une note d'introduction que laudes doit être fini avant les premières lueurs de l'aube).

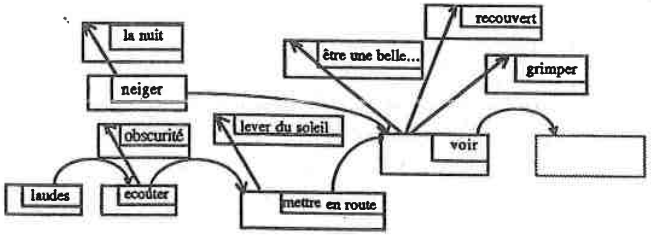
"Comme nous grimpions par le sentier abrupt qui serpentait autour du mont,..."

Nous avons ici une structure grammaticale du type "comme P(imparfait)" qui situe l'action à venir à l'intérieur de P. On génère donc une zone temporelle correspondant à "grimper par le sentier...", qui se superpose au point d'ancrage précédemment actif, puisqu'aucune indication du type plus-que-parfait ne nous en empêche (une partie a été omise) :



"...je vis l'abbaye."

Nous rencontrons pour la première fois un passé simple qui marque une suite d'actions. La zone temporelle "voir(locuteur,abbaye)" s'insère normalement au niveau de la zone active, mais contrairement à ce qui se passe dans le cas de l'imparfait, ne contient pas le nouveau point d'ancrage, elle le précède par la relation d'adjacence, indiquant le caractère fini de l'action et la possibilité de recevoir des informations sur les événements à venir. La structure finale de représentation du passage étudié est donc :



3.3.3 Quelques remarques sur le texte anglais.

First day
PRIME

In which the foot of the abbey is reached, and
William demonstrates his great acumen.

It was a beautiful morning at the end of November. During the night it had snowed, but only a little, and the earth was covered with a cool blanket no more than three fingers high. In the darkness, immediately after lauds, we heard Mass in a village in the valley. Then we set off toward the mountain, as the sun first appeared.

While we toiled up the steep path that wound around the mountain, I saw the abbey.

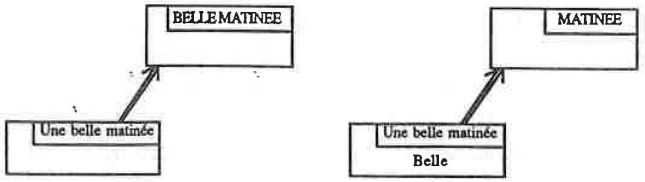
Le version anglaise donnée ci-dessus présente des différences qui posent certains problèmes de cohérence par rapport à la représentation adoptée plus haut. Elle utilise beaucoup moins les facilités que procurent le plus-que-parfait et les temps progressifs (correspondant à l'imparfait en français). En fait, l'action commence ici dès la troisième phrase ("we heard Mass"), et il semble impossible d'établir un lien direct entre les actions qui suivent et la situation exposée dans les deux premières phrases. Seul un raisonnement (sur les liens entre la nuit, le lever du soleil et l'obscurité) permet de relier deux parties de représentations que les indices linguistiques laissent indépendantes. En fait, on se trouve dans la situation générale où le lecteur (l'auditeur) doit mettre en œuvre ses facultés de raisonnement pour suivre le fil d'un discours. Le texte français est lui 'trop' explicite puisqu'il permet à lui seul de construire une représentation connexe où toute zone temporelle est reliée aux autres. Mais, même dans ce cas on constate que la représentation obtenue est très incomplète par rapport à ce qu'on attendrait intuitivement. C'est effectivement en faisant référence à des connaissances plus précises sur le monde ou le domaine, qu'il est possible de combler certains "trous", et surtout d'effectuer des prédictions à partir de ces données brutes.

3.3.4 Retour sur quelques points de détail.

Je n'ai représenté dans un premier temps la sémantique de "C'était une belle matinée de la fin novembre" que comme un bloc temporel situant le cadre de l'action. Détailler cette représentation nécessite qu'on se pose des questions plus précises sur les déterminants, le temps et l'espace, et surtout sur les classes et les instances. Je vais me limiter pour l'instant à l'exemple qui pourra servir de base à certaines généralisations.

"C'était" désigne un objet de l'espace, et l'assimile au complément qui se situe à sa droite, c'est à dire "une belle matinée de la fin novembre". Celui-ci peut s'interpréter de diverses façons suivant la représentation adoptée en mémoire. La structure "P de Q" peut être interprétée immédiatement au niveau temporel comme une relation d'inclusion. On aura donc le schéma type d'analyse (à l'entrée lexicale "de" par exemple) :

Il reste donc à interpréter P et Q dans le cas présent. "Une belle matinée" est ici la description d'une instance d'un rôle particulier (présence d'un article indéfini). On a ici le choix entre deux schémas suivant que l'on considère le rôle comme étant "belle matinée" ou "matinée", l'instance étant dans ce cas qualifiée de belle. Ceci peut se représenter ainsi :

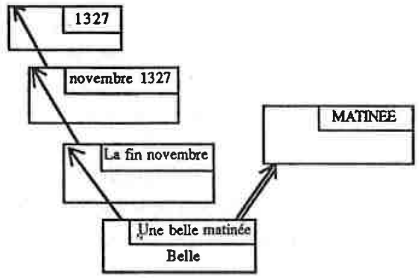


1ère possibilité 2ème possibilité

La deuxième possibilité est plus satisfaisante au niveau conceptuel (on pourrait d'ailleurs représenter la qualification "belle" comme un lien d'héritage avec le concept

"BELLE"). Beaucoup de variations sont néanmoins possibles suivant ce qui a été enregistré dans la base de connaissances.

"La fin novembre", du fait de la présence d'un article défini, représente soit un objet connu de l'interpréteur, soit un élément déductible des connaissances courantes. Ici, nous sommes dans le deuxième cas, puisque l'action (pour ceux qui ont lu le prologue!) se situe en 1327 et qu'il n'existe qu'un seul mois de novembre dans une année [connaissance du domaine]. Sans détailler l'origine de ces deux zones, on obtient la représentation suivante :



3.3.5 Conclusions partielles.

J'ai présenté ici certains éléments détaillant comment je conçois la construction temporelle d'une suite d'énoncés. Le fait de s'être attaché ici à un texte narratif n'est pas trop contraignant dans un premier temps, mais surtout, il présente l'avantage d'éluder les problèmes relatifs aux désirs et autres mondes imaginaires. C'est l'analyse de ce texte bien précis qui m'a fait prendre conscience de la nécessité d'introduire un point d'attention auquel on se réfère à tout moment de l'analyse. Ce point peut éventuellement être supplanté par un autre plus local (cf Plus-que-parfait) pour être réactivé de nouveau un peu plus tard. On sent un lien de parenté très fort avec la référence (R) dans le modèle de description des énoncés utilisé par Hornstein (77) et repris par Andrée et Mario Borillo (88). On peut reprocher à cette dernière de n'être pas assez opérative et de trop reposer sur l'intuition (On sent bien son utilité, mais on ne voit pas comment l'obtenir ni comment l'utiliser). Au contraire, il est possible d'établir des règles permettant de créer les points d'ancrage, et surtout de les faire évoluer, au fur et à mesure de l'analyse d'énoncés.

3.4 Objets et mécanismes complémentaires.

3.4.1 Les zones de cohérence.

1 Introduction.

L'exemple précédent a permis de montrer les principaux mécanismes qui conduisent à l'interprétation puis à la compréhension d'une série d'énoncés dans le cadre d'un texte narratif. Alors que nous avons vu comment une phrase pouvait s'insérer dans le contexte généré par celles qui l'ont précédée, nous n'avons pas explicité comment nous associons à une portion d'énoncé une représentation interne correspondant à son contenu. Le concept de *zone de cohérence* est justement l'outil qui intègre cette association dans le cadre de notre modèle. Cet outil se révèle bien plus puissant encore car il permet de généraliser la notion d'association entre concepts.

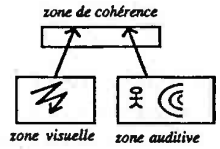
Nous développerons ensuite un autre point important : le traitement du présent français qui a déjà posé pas mal de problèmes à ceux qui voulaient l'aborder sous l'angle "ERS" de Hornstein. La zone de cohérence nous donne un moyen élégant de le représenter tout en autorisant les multiples valeurs qu'il est susceptible d'avoir.

2 Comprendre un message.

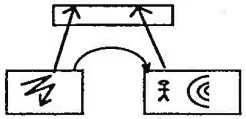
Il est parfois des situations où, pour deux événements particuliers, nous ne sommes pas capables, dans une première analyse, de déterminer leurs positions relatives. Inclusion ou Adjacence ? La réponse de Allen à cette difficulté est de marquer entre les deux intervalles concernés l'ensemble des relations qui sont admissibles. Cette méthode pose un problème : elle ne distingue pas l'association privilégiée qui existe entre deux événements. Ainsi, si pendant un orage on perçoit un éclair, puis, quelques instants après, on entend le grondement de la foudre, quel que soit le temps qui les sépare, nos habitudes nous les font percevoir dans une même cohérence. En particulier, si ces deux événements sont très rapprochés (l'orage aussi), nous pouvons à la limite les percevoir comme un seul événement dont les facettes sont multiples.

Pour représenter cette association spécifique entre plusieurs zones perçues ou construites, nous introduisons la notion de *zone de cohérence*, zone similaire aux autres, mais qui contient (au sens de l'inclusion toujours) un certain nombre d'autres zones qui

de ce fait sont liées entre elles. Reprenons l'exemple de l'orage, les deux perceptions, visuelles et auditives, peuvent être représentées de la façon suivante :



Dans le cas particulier où l'on dispose d'informations pour situer les deux événements, ceux-ci peuvent être liés par une relation d'adjacence comme ci-dessous :



Si nous considérons les schémas précédents de façon asymétrique, il est possible d'envisager une des deux zones liées, comme une clef d'accès à la structure complète. Ainsi, si on perçoit un éclair, l'*instanciation* d'une zone de cohérence similaire à celle ci-dessous va nous permettre de prédire l'apparition d'un grondement de foudre dans les instants à venir. Inversement, si nous entendons le tonnerre, nous pouvons supposer qu'il y a eu un éclair auparavant. Ce mécanisme d'accès et de *prédiction* fonctionne donc dans les deux sens, et plus généralement, l'accès à toute structure complexe telle que celle représentée figure 3.4 peut se faire à partir de n'importe quel de ses sous-éléments.

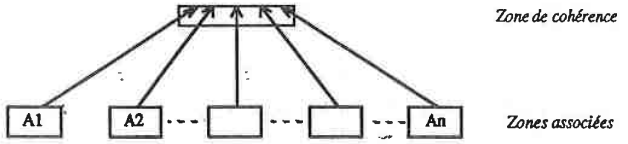


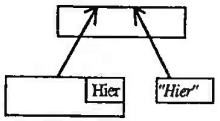
Figure 3.4 : Schéma général d'une zone de cohérence.

L'origine des différentes zones qui se trouvent associées sous une même zone de cohérence peut être très variable. Il peut s'agir d'un ensemble de perceptions associées à un même phénomène, par exemple un bruit, un aspect et une odeur d'essence pour un moteur, mais aussi d'un *signe qui représente* un objet ou une action. Ainsi, à la zone correspondant à la perception d'un chat, on peut envisager d'associer le mot "chat" et

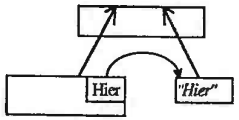
diverses autres informations relatives à ce concept. Ce que nous appelons alors concept est justement *la* zone de cohérence qui relie toutes ces informations.

3 RETOUR AU LANGAGE.

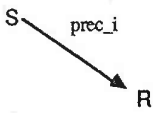
Dans le paragraphe précédent nous venons d'entrevoir un schéma d'analyse de message qui conduit à l'intégration dans le même espace de représentation des éléments de la langue et de leurs significations. En effet, chaque partie d'un énoncé exprimant quelque chose au niveau temporel va s'insérer dans une structure de cohérence qui contient cette représentation. Par exemple, au mot "hier" , on peut associer une zone représentant le temps correspondant à celui-ci quand le locuteur le prononce :



Le schéma ci-dessus représente l'interprétation initiale qui peut être faite de "hier" par un auditeur ne connaissant pas la valeur temporelle de ce mot. Cependant, si celui-ci possède des informations supplémentaires, comme par exemple le fait que *'hier'* (le concept) se déroule toujours avant son énonciation, on peut obtenir une structure plus complexe :

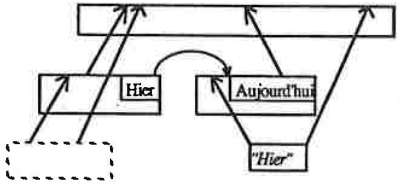


On peut alors rapprocher cette représentation de celle proposée par A. et M. Borillo (88) dans le formalisme "ERS" qui situe l'énoncé et l'intervalle correspondant grâce aux relations de Allen (on se référera aussi à [Yip 85]) :

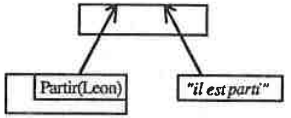


Référence antérieure à l'intervalle de discours.

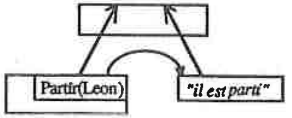
Cependant, notre méthode permet d'introduire la notion de présupposition quand les connaissances de l'auditeur sont encore plus importantes, et font notamment intervenir une référence au jour courant de l'énonciation dans lequel celle-ci est incluse et qui succède à la zone indiquée par "hier". Si de plus, on sait que "hier" n'indique qu'une référence temporelle dans laquelle doit se situer une action que l'on va décrire, on obtient la représentation plus complète suivante (elle n'est pas la seule possible) :



La création d'une représentation ne se fait pas de la sorte uniquement au niveau du mot. Elle peut se faire pour des syntagmes, phrases, dialogues ou textes, pour peu que ceux-ci relatent une période homogène. Dans le cas d'un énoncé simple, l'association est souvent due au prédicat employé qui exprime une situation se déroulant dans le temps. Pour l'énoncé "Leon est parti", on a ainsi la construction de la structure suivante :



Là encore, si l'auditeur de cet énoncé tient compte du temps de conjugaison, il pourra construire une représentation où l'action sera antérieure à l'énoncé comme suit :



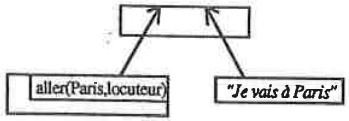
Cette manière d'interpréter les expressions de la langue permet de considérer sous un nouvel angle l'exemple du *Nom de la Rose* traité au 3.3. Tout en conservant la même valeur pour les différents temps de conjugaison que nous avions analysés, nous

comprenons maintenant comment le lien est obtenu entre l'énoncé perçu et la construction qui en est faite. Il suffit d'introduire à chaque fois la zone de cohérence qui s'impose.

Parmi les temps de conjugaison qui ont servi d'exemple jusqu'ici, on a surtout trouvé les passés (simple, composé et imparfait) et nous n'avons rien dit des autres temps du français qui doivent, au même titre que ces derniers, être considérés par notre modèle. Nous allons ici nous intéresser au présent qui possède un comportement particulier au niveau temporel, pour ne traiter les temps restants que plus tard lors d'une description plus complète des rapports entre le temps et le langage.

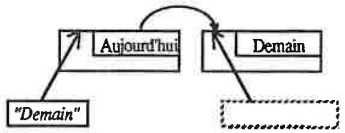
Il est classiquement reconnu que le présent conduit à de nombreuses interprétations temporelles en fonction de son contexte d'analyse. Comme le remarque F. Nef (84), il peut exprimer un passé proche dans le cas d'un commentaire journalistique : "Platini marque un but - non ! la balle rebondit sur la barre, mais il la reprend de volée - oui ! le but est marqué!" (p.84) ; parfois le temps de la situation contient fréquemment le temps de l'énoncé : "le ciel est bleu" ou il peut avoir une valeur de futur plus ou moins proche : "je pars" (Nous n'avons pas encore tenu compte des valeurs générales ou répétitives du présent pour l'instant, bien que nous ayons conscience de l'importance du problème). Il est donc difficile d'associer à coup sûr une relation temporelle entre l'énoncé et la situation qu'il exprime sans crainte de voir celle-ci remise en question.

Prenons l'exemple de la phrase "Je vais à Paris". Nous ne pouvons, sans connaître le contexte, déterminer si elle exprime un vrai présent (sur le quai d'une gare) ou un futur lointain (prévision pour des vacances). Par contre nous savons qu'elle décrit une situation dans laquelle *je vais à Paris* quelle que soit la période où ceci se déroule. Nous pouvons donc construire la représentation suivante :

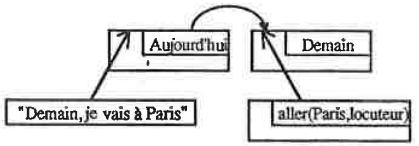


Et c'est tout ce que nous pouvons tirer de l'énoncé, c'est à dire des seules informations linguistiques. Nous resterons donc sur cette représentation *intemporelle* du présent français qui laisse toute liberté à des informations contextuelles pour compléter les schémas de représentation obtenus. ***Le présent est une simple association d'un message à son contenu par l'intermédiaire d'une zone de cohérence.***

Comment va alors intervenir le contexte de la phrase pour construire une représentation temporelle complète? Il peut s'agir d'informations générales sur le domaine qui permettent de décomposer la situation de telle manière que l'on puisse situer systématiquement le temps de l'énonciation par rapport à celle-ci ; mais dans les cas ambigus comme "Je vais à Paris", le syntagme verbal est souvent accompagné de locutions qui précisent ce lien temporel. Si nous considérons l'énoncé : "Demain, je vais à paris", en associant à 'demain' la construction suivante (on a omis la zone de cohérence par souci de clarté) :



La zone associée à l'action décrite par "je vais à Paris" se situe alors naturellement dans le point d'attention laissé en suspend, pour donner :



Une rapide déduction temporelle (cf 3.5.2) permet alors de déduire que cette action est située dans le futur par rapport à l'énoncé (les deux zones sont incluses respectivement dans deux zones adjacentes).

4. ORIGINE DES ZONES DE COHÉRENCE.

Pour les différentes constructions que nous avons réalisées à partir de portions d'énoncés, nous avons indiqué que plusieurs possibilités existaient en fonction du degré de connaissance du système (ou de l'auditeur en général) qui construit cette représentation. Toute zone de cohérence est donc, dans notre approche, apprise à partir d'observations antérieures. Nous n'envisageons pas, bien évidemment de proposer une théorie complète de l'apprentissage de la langue, qui dépasserait d'ailleurs nos compétences linguistiques et psychologiques. Néanmoins, on peut se demander si les zones de cohérence ne pourraient pas servir de fondement à une telle analyse.

Dans le cas d'un système de dialogue qui se veut rapidement opérationnel, il sera nécessaire d'initialiser la base de connaissance grâce à un ensemble de schémas standards de cohérence couramment rencontrés dans la langue ou dans l'univers de l'application. Pourtant, rien n'interdira au système, sur la base de ces schémas, de construire de nouvelles représentations, en particulier quand l'usager utilisera des formes qui ne lui seront pas connues.

5 REMARQUES DIVERSES ET CONCLUSIONS.

Nous avons introduit dans cette partie un élément important de notre modèle : les zones de cohérence. Bien qu'il s'agisse d'un schéma relativement général d'analyse, elles nous sont apparues nécessaires pour étudier plus spécialement les mécanismes de compréhension d'un énoncé dans un contexte quelconque. Malgré la spécificité apparente de ce concept, il ne nous est pas apparu nécessaire de distinguer explicitement les zones de cohérence des autres zones de notre modèle. Au contraire, le principe de base de celui-ci est qu'une zone particulière est un ensemble d'informations cohérentes à un niveau inférieur et donc toute zone est potentiellement une zone de cohérence.

Deux points particuliers liés aux zones de cohérence n'ont pas été abordés, à savoir comment celles-ci interviennent dans l'étude des structures "syntaxiques" de la langue et dans les mécanismes de raisonnement faisant intervenir la notion de schémas standards. Nous reviendrons sur tout ceci au 3.6.

Pour l'instant, nous avons entre nos mains un outil qui semble relativement puissant pour les quelques exemples que nous avons traités. Cependant, en l'état actuel de nos recherches, nous ne le maîtrisons pas encore totalement et de nombreux problèmes à la fois généraux et de détails restent à régler.

3.4.2 Nécessité d'un mécanisme d'héritage.

1 INTRODUCTION.

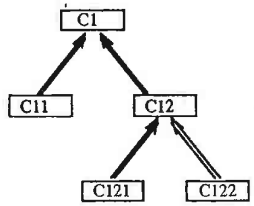
Durant l'analyse de l'exemple du *Nom de la rose*, nous avons rencontré maintes fois la nécessité de posséder des schémas standards pouvant être utilisés à tout moment d'une analyse linguistique ou d'un raisonnement. En effet, reconnaître un énoncé, c'est retrouver parmi les mots qui le composent des structures ou au moins des parties de

structure déjà connues du système. De même, raisonner sur un univers, c'est reconnaître que telle information correspond à une prémisse d'un schéma de déduction préalablement établi. En fait, les mécanismes d'apprentissage et de prédiction dont nous avons déjà parlé passent par la maîtrise de la construction et de la réutilisation de ces schémas.

L'utilisation de mécanismes d'héritage correspond assez bien à ce genre de manipulations. En effet, il permet d'attacher certaines propriétés à une classe d'objets ; de sorte que tous les éléments ou instances de cette classe en héritent quand cela est nécessaire.

2 PROPRIÉTÉS GÉNÉRALES.

Les notions d'héritage et de typicalité peuvent conduire à des développements relativement complexes (notamment dans le cadre des langages orientés objets) que nous n'aborderons pas ici. Nous nous appuierons sur le schéma de base suivant où chaque classe possède un certain nombre d'instances, qui à leur tour peuvent former autant de classes (ce paradigme est différent de celui utilisé en Flavor, où les classes et les instances forment deux ensembles distincts) :

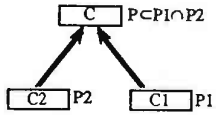


A partir de ce schéma deux grands choix sont possibles pour la gestion des propriétés le long de l'arborescence ainsi créée par la relation d'héritage :

- les différentes propriétés de classe sont groupées au niveau du nœud de cette classe, de sorte que chacune des instances de celle-ci va hériter de toutes les propriétés de son ancêtre. Dans le cas contraire, l'instance se comporte comme une exception et la mise en œuvre de l'ensemble doit comporter des mécanisme de gestion des conflits qui peuvent résulter de cette situation ([Touretzky 86]).
- l'autre solution consiste à ne remonter au niveau d'une classe que les propriétés vraiment communes à toutes ses instances. Une grande partie du comportement de la classe résulte alors des particularités de chacun de ses éléments, ce qui donne une très

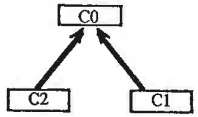
grande souplesse de fonctionnement à l'ensemble. Nous avons adopté cette option, car elle correspond le mieux aux types de traitement que nous désirons effectuer le long de la hiérarchie des classes.

Si par exemple, deux classes C1 et C2, dont les ensembles de propriétés sont respectivement P1 et P2, héritent tous deux d'une classe C, dont l'ensemble des propriétés est P, on aura $P1 \cap P2 \supset P$. Ceci se schématise ainsi :

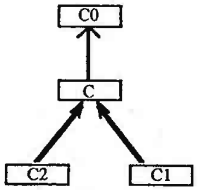


Cela signifie que toute classe est effectivement l'objet le plus général émanant d'un ensemble d'instances. Dans ces conditions deux opérations fondamentales peuvent être effectuées le long de la relation d'héritage : l'*abstraction* et l'*instanciation*.

L'abstraction consiste à créer une nouvelle classe plus générale à partir de deux (ou plusieurs) classes existantes. Soient donc deux classes C1 et C2 de propriétés respectives P1 et P2, créer une abstraction de C1 et C2 revient à générer une classe C dont les propriétés P sont exactement égales à $P1 \cap P2$. De la sorte, on préserve les contraintes exprimées ci-dessus entre une classe et ses instances et d'autre part, on maximalise l'ensemble des propriétés supportées par C. Il se peut aussi que C1 et C2 héritent déjà d'une classe générale C0 comme dans la figure ci-dessous :

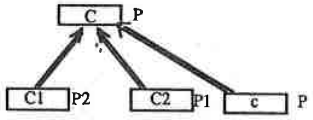


Pour préserver la cohérence de la relation d'héritage, il est alors nécessaire que la nouvelle classe C hérite elle aussi de cette classe C0 de la façon suivante :



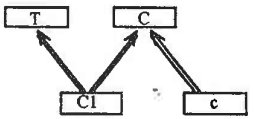
L'abstraction à partir de deux objets doit donc à la fois préserver leurs propriétés communes, mais aussi les liens d'héritage qu'ils partagent.

L'opération inverse de l'abstraction : l'instanciation consiste à compléter l'ensemble des propriétés d'un objet donné que l'on désire mieux connaître. Soit donc une classe C dont les propriétés sont P et dont on connait déjà une instance C1 de propriétés P1 et C2 de propriétés P2. Supposons que l'on crée une nouvelle instance c de C. Par défaut et suivant les propriétés de notre relation d'héritage, c va déjà hériter de toutes les propriétés de C. On aura donc :

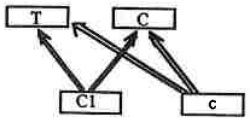


Pour connaître alors plus de propriétés sur c, il faut aller en chercher parmi les autres instances de la même classe C, c'est à dire ici les propriétés de C1 ou de C2. Afin de conserver une régularité de fonctionnement à l'ensemble, les nouvelles propriétés de c seront par exemple $(P \cap P1)$, ou $(P \cap P2)$ ou $(P \cap P1 \cap P2)$ suivant les choix stratégiques d'instanciations qui auront été faits.

Là encore peuvent intervenir d'autres relations d'héritage qu'établissent éventuellement les anciennes instances C1 ou C2. Ainsi si on se trouve dans la situation suivante, où C1 hérite d'une classe T en plus de C :



un apport d'information (de propriétés) au niveau de c peut consister alors à faire aussi hériter c de T, en transmettant au passage les propriétés correspondantes vérifiées par T. On a alors :

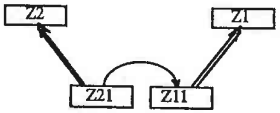


Ce dernier phénomène complète alors l'éventail des opérations qui peuvent être effectuées sur le lien d'héritage entre deux classes.

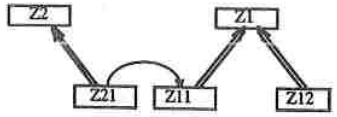
3 APPLICATION AUX ZONES TEMPORELLES.

Dans le cas de notre modèle, les seules propriétés à considérer sont les relations temporelles qu'une zone établit avec d'autres. La relation d'héritage que nous désirons mettre en œuvre va donc porter sur les seules zones, avec propagations des liens temporels d'une zone à une autre de la même classe.

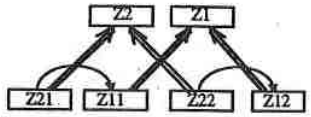
Etudions le fonctionnement du mécanisme d'héritage sur un exemple. Supposons que l'on ait une zone Z1 dont une instance Z11 établisse une relation d'adjacence avec une instance Z21 d'une classe (zone) Z2 comme sur le schéma ci-dessous.



Si maintenant, une nouvelle zone apparaît, dont on sait qu'elle hérite directement de Z1 et dont on ne connaît aucune des propriétés :



Il est alors possible d'induire certaines propriétés temporelles sur Z12, soit la liant directement à Z21, soit en instanciant une nouvelle zone descendant de Z2 de la façon suivante :



Le choix entre ces deux solutions peut résulter d'un principe circonscriptif ([Kayser 88] d'après [McCarthy 86]) dont l'usage ne doit pas être systématique et dépend explicitement des circonstances. Nous ne développons pas plus pour l'instant les propriétés de la relation d'héritage sur les zones temporelles. Dans l'ensemble elles sont tout à fait similaires aux schémas généraux que nous avons présentés dans la section précédente. Cependant des conséquences plus précises en rapport avec notre modèle seront exposées ultérieurement au chapitre 3.5.3.

4 CONCLUSIONS.

Nous disposons maintenant de l'essentiel des outils de notre modèle de représentation des informations temporelles. En plus des zones et des relations d'adjacence et d'inclusion, nous venons d'introduire une relation d'héritage qui va servir de base à tous les mécanismes d'abstraction et d'instanciation. Pourtant, il faut bien distinguer deux aspects bien différents de notre modèle. D'un coté nous disposons d'outils de base (zones et relations) qui forment un matériau fixe et de l'autre des éléments plus dynamiques dont les zones de cohérences sont un bon exemple. En particulier, les problèmes de stratégie, qui à ce stade de notre réflexion ne sont pas encore complètement résolus, gardent une importance primordiale sur la qualité du comportement de notre modèle.

3.5 Les opérations à mettre en œuvre.

3.5.1 Introduction.

Après avoir introduit les éléments principaux manipulés dans le modèle de représentation proposé, il apparaît nécessaire de pouvoir effectuer certaines opérations sur celui-ci, afin qu'il ne soit pas qu'un simple instrument descriptif, mais qu'il puisse aussi induire des processus de déduction et d'apprentissage que possède tout système dit intelligent.

Nous pouvons dans un premier temps remarquer deux grandes catégories d'opérations possibles sur nos objets, en fonction des deux types de relations existant entre ceux-ci. Les relations d'inclusion et d'adjacence impliquent un raisonnement local pour vérifier la cohérence, et éventuellement pour étendre ces relations pour un ensemble de zones déterminé. Ces déductions sont alors purement temporelles et peuvent donc être vues comme relativement horizontales au niveau du système tout entier.

Une autre catégorie de raisonnements correspond au parcours des relations d'héritage entre différents niveaux de zones temporelles. Suivant un parcours ascendant ou descendant le long de ces relations, il est possible de faire soit des opérations d'inductions sur des zones existantes, soit de la déduction par instanciation d'éléments inconnus. Ces deux opérations vont permettre de créer des zones particulières qui vont restructurer partiellement la hiérarchie. Nous étudierons alors ces mécanismes sur la base des principes généraux que nous avons établis dans la section précédente, notamment en ce qui concerne les opérations d'intersection et de réunion d'ensembles de propriétés.

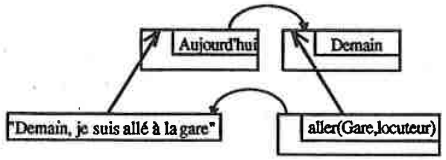
3.5.2 Les déductions temporelles.

1 Introduction.

Puisque nous manipulons des relations temporelles entre des zones, nous devons nous assurer que l'ensemble forme continuellement un tout cohérent. En effet, notre modèle de représentation tout entier repose sur ces zones et une erreur qui s'y glisserait pourrait perturber fortement son fonctionnement. Cependant, comme les données à manipuler sont très nombreuses et variées, vérifier cette cohérence sur l'ensemble de la base de connaissance risque de monopoliser toutes les ressources du système. C'est pourquoi

nous n'adopterons pas un calcul exhaustif, mais plutôt une méthode rapide et qui donne malgré tout de bons résultats.

Considérons la phrase : "Demain, je suis allé à la gare". C'est un exemple typique où il est nécessaire de faire intervenir le mécanisme de vérification de cohérence avant de valider la construction d'une représentation. En effet, la structure naturelle qui serait obtenue à partir de cette phrase est :



Un système de raisonnement temporel doit au minimum détecter ces cas d'incohérence en propageant les contraintes associées aux différentes relations au sein de l'ensemble des zones.

La vérification de la cohérence n'est pas obligatoirement la seule tâche allouée à un raisonneur temporel, il peut aussi compléter les liens temporels qui se déduisent des informations existantes dans la base. Ces deux opérations de vérification de la cohérence et de complétion des relations correspondent en fait à un seul calcul dont on utilise différemment les résultats.

2 Méthode adoptée.

La méthode adoptée par Allen pour vérifier la cohérence des relations existant entre des intervalles consiste à propager des contraintes le long du graphe ainsi formé. Cette opération repose sur une table des transitions élémentaires déduites à partir de deux relations. On trouvera cette table et l'algorithme de déduction dans [Allen 83].

Contrairement à Allen qui manipulait en tout treize relations, nous n'avons plus à gérer que 4 relations (deux plus leurs inverses). La difficulté du calcul, presque rédhibitoire dans le cas de Allen, devient pour nous beaucoup plus simple. Tout d'abord, comme cela apparaît table 3.1, le nombre des transitions élémentaires est fortement réduit et surtout, une seule relation est obtenue dans la plupart des cas. Pour trois zones temporelles a, b et c et deux relations r_1 et r_2 telles que $(a\ r_1\ b)$ et $(b\ r_2\ c)$, nous avons indiqué les relations

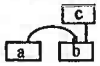
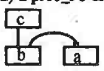
pouvant exister entre a et c. Dans trois cas ("*pas d'info*"), aucune contrainte n'apparaît entre a et c, toutes les relations sont donc admissibles entre ces deux zones. Dans cinq cas (numérotés dans la table 3.1), une ambiguïté subsiste, puisque deux relations parmi quatre sont autorisées sans que l'on dispose de plus d'information. Enfin dans huit cas sur seize, une relation est immédiatement déductible entre a et c.

a r1 b \ b r2 c	prec	prec_i	in	in_i
prec	*prec*	pas d'info	{prec,in} (1)	*prec*
prec_i	pas d'info	*prec_i*	{prec_i,in} (2)	*prec_i*
in	*prec*	*prec_i*	*in*	pas d'info
in_i	{prec,in_i} (3)	{prec_i,in_i} (4)	{in,in_i} (5)	*in_i*

Table 3.1 : transitions élémentaires pour deux relations r_1 et r_2.

Comment une incohérence est elle décelée dans ces conditions ? Pour chaque déduction directe qui peut être faite à partir de la table 3.1 (les huit cas) si la relation induite entre a et c est différente d'une relation déjà existante, il y a contradiction, puisque entre deux zones ne peut être établie qu'*une seule relation temporelle*. Pour les cinq cas ambigus, il serait aussi possible de déduire un incohérence par rapport à une relation déjà existante, mais cette opération se révèle inutile comme nous allons le voir.

En effet, une propriété particulière ressort du faible nombre de relations que nous manipulons. Si une relation interdite apparaît entre deux zones où une transition ambiguë a été calculée, les déductions effectuées à partir de cette relation génèrent automatiquement une incohérence. Nous montrons cette propriété par l'intermédiaire de la table 3.2 qui présente l'ensemble des cas possibles entre 3 zones a, b et c. Pour chaque couple de relations (r1, r2) tel que {a r1 b} et {b r2 c}formant une transition ambiguë (cas 1 à 5 de la table 3.1), la table 3.2 indique pour chaque relation qui se trouve interdite, l'inférence qui serait possible et la contradiction qui en résulterait.

Transition ambiguë et schéma de la situation	Arc interdit	Inférence possible	Contradiction
1) a prec b & b in c 	a prec_i c	a prec_i c & c in_i b	a prec_i b ≠ a prec b
	a in_i c	a in_i c & c in_i b	a in_i b ≠ a prec b
2) a prec_i b & b in c 	a prec c	a prec c & c in_i b	a prec b ≠ a prec_i b
	a in_i c	a in_i c & c in_i b	a in_i b ≠ a prec_i b

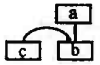
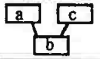
3) a in_i b & b prec c 	a prec_i c a in c	a prec_i c & c prec_i b b in a & a in c	a prec_i b ≠ a in_i b b in c ≠ b prec c
4) a in_i b & b prec_i c 	a prec c a in c	a prec c & c prec b b in a & a in c	a prec b ≠ a in_i b b in c ≠ b prec_i c
5) a in_i b & b in c 	a prec c a prec_i c	b in a & a prec c b in a & a prec_i c	b prec c ≠ b in c b prec_i c ≠ b in c

Table 3.2 : Interdictions résultant des ambiguïtés de la table 3.1.

La conséquence de cette propriété est double :

- Pour les déductions ambiguës, la vérification de la cohérence est inutile, car dans le cas où il existe une incohérence, celle-ci se transforme en une déduction simple qui s'oppose aux relations initiales.

- Nous ne devons pas, comme Allen, propager des ensembles de contraintes entre les zones temporelles. Seules les déductions simples sont utilisées dans la table 3.1.

Nous pouvons maintenant, présenter l'algorithme de vérification de cohérence parmi un ensemble de zones $\mathcal{Z}$. Soit $\mathcal{A}$ les arcs connus du système liant les zones de $\mathcal{Z}$ et $\mathcal{New\_A}$ l'ensemble des arcs calculés à l'issu de l'algorithme. On utilisera la fonction **transition**(k_1,k_2) qui calcule la déduction résultant des deux arcs k_1 et k_2, conformément à la table 3.1 et **contradictoire**(k,Arcs) qui regarde si un arc de Arcs ne s'oppose pas à k (ne concerne pas les même zones avec une relation différente). La procédure filtre aussi les redondances éventuelles parmi l'ensemble d'arcs par la fonction : **existe**(k,Arcs).

```
New_A = A ;
Pour rel1(zi , zj) dans New_A ;                 /* une relation rel1 liant zi et zj */
   Debut ;
   New_A = New_A - {rel1(zi , zj)} ;
   si contradictoire(rel1(zi,zj) , New_A ) ;
         RETOURNER ERREUR ;                     /* détection d'une incohérence */
   sinon si existe(rel1(zi,zj) , New_A ) ;
         continuer ;                            /* on étudie l'arc suivant */
   sinon
         Debut ;
         soit rel2(zi,zk) et rel3(zk,zj) deux arcs de New_A ;
         X = transition(rel1(zi,zj),rel2(zi,zk)) ;
         si cardinal(X) = 1 ;                   /* une déduction simple */
               New_A = New_A + X ; /* on met à jour les arcs */
         Fin ;
```

```
Fin ;                          /* la procédure s'arrête quand il n'y a plus d'arcs à explorer */
RETOURNER New_A ;
```

3 CONCLUSION.

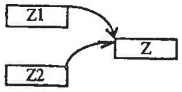
Les mécanismes de vérification de la cohérence temporelle semblent particulièrement simples dans notre modèle de représentation. Pourtant, les temps de calcul risquent de devenir rédhibitoires dès que la quantité de zones manipulées devient vraiment trop importante. c'est pourquoi il faut envisager, comme le fait Allen (83), de limiter le raisonnement à certains sous ensembles de la totalité des zones temporelles. A ce titre, les zones de cohérence représentent un outil parfaitement adapté pour délimiter des espaces séparés de raisonnement.

3.5.3 Les raisonnements liés à la hiérarchie.

1 INTRODUCTION.

Nous avons présenté au chapitre 3.4.2 le mode de fonctionnement pour le mécanisme d'héritage que nous avons décidé d'adopter, qui repose sur deux opérations fondamentales : l'instanciation et l'abstraction. Ces opérations, qui correspondent à l'échelle de tout un système à des phénomènes d'*apprentissage* et de *prédiction*, ont été à cet endroit rapidement appliquées aux zones temporelles : il est maintenant temps de préciser la nature exacte des mécanismes qui, dans notre modèle, leur sont associés.

Comme nous l'avions remarqué, les propriétés manipulées le long de la relation d'héritage se limitent pour les zones temporelles aux relations que celles-ci établissent avec d'autres zones. Ceci implique que nous puissions définir pour ces relations les deux opérations ensemblistes d'*intersection* et d'*union*. Or, si nous désirons que ces opérations ne soient pas trop contraignantes, il n'est pas conseillé de définir une propriété commune à deux zones Z1 et Z2 comme étant l'établissement d'une relation de *même type* avec une *même troisième zone* Z, comme dans la figure ci-dessous :



Il est plus intéressant d'envisager l'égalité de deux propriétés comme étant l'établissement d'une relation de même type avec deux nouvelles zones de *même nature*,

c'est à dire possédant au minimum un ancêtre commun - le cas de la figure précédente pouvant être inclus dans notre définition.

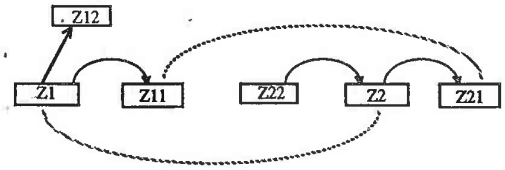
Une telle hypothèse rend beaucoup plus complexe l'évaluation des propriétés communes ou différentes entre deux zones, dans le but de calculer une intersection ou une union. Il faut pour ceci mettre en œuvre un mécanisme d'appariement qui, pour deux zones données, propose une mise en correspondance possible entre elles, sachant que le résultat obtenu a peu de chance d'être unique.

Nous commencerons donc par analyser une procédure possible d'appariement entre deux zones, pour ensuite montrer la réalisation effective des mécanismes d'abstraction et d'instanciation. Il est remarquable d'observer que les traitements mis en œuvre sont en fait très proches de ceux envisagés par B. Moulin (88) pour les graphes conceptuels (généralisation et jointure). Il serait intéressant d'étudier jusqu'à quel point cette analogie correspond à une même manière de modéliser certains phénomènes.

Enfin, nous proposerons l'étude d'un exemple comme application du mécanisme d'instanciation à la reconnaissance d'une phrase complexe, en introduisant les prédictions comme des cas particuliers d'instanciation.

2 APPARIEMENT DE ZONES.

Apparier deux zones, ainsi que leurs voisins respectifs (au sens des relations d'adjacence et d'inclusion), c'est mettre en correspondance ces deux zones, ainsi que les relations du même type. Par exemple, la figure suivante montre deux zones Z1 et Z2 et un appariement possible à partir de celles-ci (indiqué en pointillé).

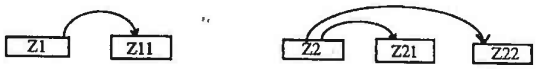


Nous notons ainsi l'appariement précédent : {(Z1, Z2), (Z11, Z21), (Z12, $vide), ($vide, Z22)} où '$vide' indique que l'élément en vis à vis n'a pas été apparié. On remarque que le couple ($vide, $vide) n'a alors pas de sens dans notre optique.

Dans le cas où les relations au départ des deux zones étudiées sont plus nombreuses et forment deux graphes complexes, l'appariement impose un parcour simultané de ceux-ci, relation par relation. Dans un premier temps, nous pouvons donner les conditions pour que deux zones ou relations puissent être appariées.

Nous avons déjà vu que pour deux zones temporelles, la condition minimale d'appariement est qu'elles possèdent des ancêtres communs. Cette condition ne peut être réduite à des ancêtres directes (un seul lien d'héritage), puisque les opérations d'abstraction que nous avons envisagées sont susceptibles de générer des nœuds intermédiaires entre deux zones. Il n'y aurait pas alors de régularité de comportement de l'opération d'appariement.

Dans le cas des relations, la seule restriction est qu'elles soient de même nature (deux inclusions ou deux adjacences) et bien sûr parcourues dans le même sens pendant l'opération d'appariement. Les contraintes à ce niveau sont donc relativement faibles ce qui autorise en particulier l'apparition d'ambiguïtés comme pour les zones Z1 et Z2 suivantes :



Les deux appariements {(Z1, Z2), (Z11, Z21), ($vide, Z22)} et {(Z1, Z2), (Z11, Z22), ($vide, Z21)}sont possibles, bien plus, on assiste à des cas extrêmes tels que {(Z1, Z2), (Z11, $vide), ($vide, Z21), ($vide, Z22)}. Les deux zones initiatrices de l'appariement restent dans tous les cas associées, car c'est cette association qui sert de base à l'étude.

L'algorithme d'appariement (au demeurant relativement long) ne sera pas décrit ici. Nous pouvons simplement signaler ici qu'il prend en entrée deux zones temporelles, et fournit comme résultat, l'*ensemble* des appariements possibles à partir de celles-ci, en tenant compte de toutes les contraintes énoncées plus haut. Un problème important apparaît alors : quel appariement choisir parmi tous ceux qui sont obtenus ? La réponse à cette question dépend d'options stratégiques qui auront été choisies dans le cadre d'un système réel et qui peuvent être très variables. Au niveau de description théorique où nous nous situons pour l'instant, nous supposerons avoir choisi systématiquement le "bon" appariement pour l'exemple traité. Nous reconnaissons que cette position est plus que discutable, mais c'est pour nous la seule façon de progresser dans nos réflexions.

3 ABSTRACTION ET INSTANCIATION DE ZONES.

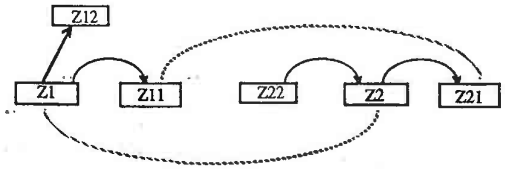
Dès que l'on possède pour deux zones données Z1 et Z2 un appariement possible $\mathcal{A}$, il est possible de passer à une phase d'instanciation ou d'abstraction conformément à l'alternative présentée au chapitre 3.4.2.

Le mécanisme d'abstraction passe par un calcul de l'intersection des propriétés communes à Z1 et Z2. En termes de relations temporelles, nous ne conservons donc que les couples complets, c'est à dire ceux dont aucun des éléments n'est égal à '$vide'. Pour chacun de ces couples (Z1i, Z2j), nous avons vu que la condition d'appariement était l'existence d'ancêtres communs à Z1i et Z2j. L'abstraction consiste alors à faire hériter ces deux éléments d'une nouvelle zone dont les propriétés sont les suivantes :

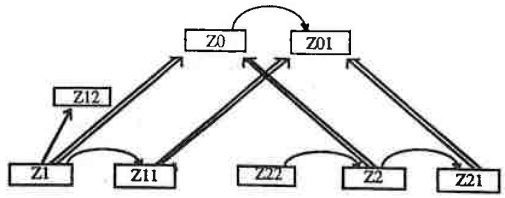
- elle hérite, elle aussi, des ancêtres communs à Z1i et Z2j, ce qui assure la cohérence de comportement de la nouvelle zone par rapport au mécanisme d'héritage.

- elle établit avec les autres zones d'abstraction créées, les mêmes liens temporels que ceux établis par Z1i et Z2j avec leurs voisins respectifs (et appariés). On rappelle que Z1i et Z2j établissent en fait les mêmes relations, d'après le fonctionnement du mécanisme d'appariement.

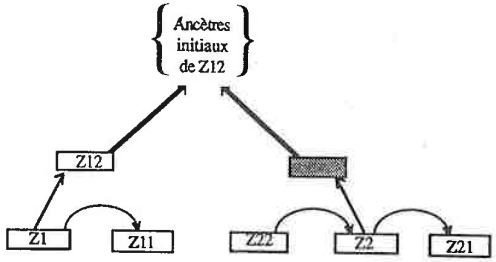
Si nous reprenons l'exemple donné pour l'appariement {(Z1, Z2), (Z11, Z21), (Z12, $vide), ($vide, Z22)} :



nous pouvons schématiser la préservation des propriétés temporelles par la figure ci-dessous, où les ancêtres communs à Z11 et Z21 (et donc à Z01 dans la figure) n'apparaissent pas par souci de clarté :



Inversement, le mécanisme de prédiction ne s'intéresse qu'aux appariements incomplets. Par exemple, si on désire compléter les propriétés attachées à une zone Z2 par comparaison avec une zone Z1 et que l'appariement de Z1 et de Z2 donne l'ensemble $\mathcal{A}$; on ne gardera que les couples de $\mathcal{A}$ dont le *deuxième* élément est vide (ce qui correspond à une propriété nouvelle pour Z2). On génère pour chacun de ces couples une nouvelle zone possédant exactement les mêmes propriétés (temporelles et hiérarchiques) que leurs membres gauches (rappel : le mécanisme d'instanciation, contrairement à celui d'abstraction est asymétrique par rapport aux deux zones étudiées). Si nous reprenons de nouveau le même exemple, en supposant que c'est Z2 dont nous voulons connaître plus de propriétés, nous obtenons : (la zone grisée est celle qui a été instanciée)

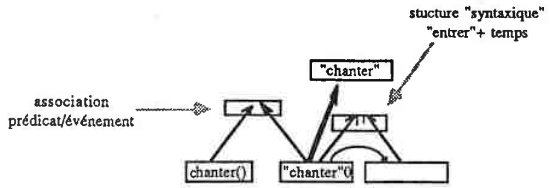


4 EXEMPLE.

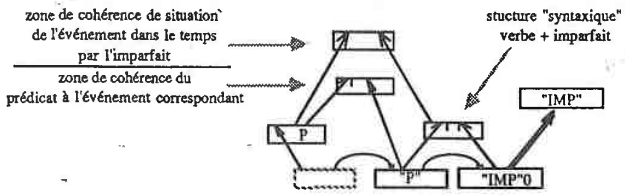
Considérons l'énoncé : "Paul chantait quand Jean est entré" que nous traduirons, dans un langage simplifié nécessaire pour la clarté de l'analyse en : "chanter IMP quand entrer PC". Nous considérons ici que la marque du temps de conjugaison est systématiquement détachée de son verbe, et située à sa droite pour des facilités d'analyse. Nous allons raisonner au niveau structurel, afin de montrer comment deux schémas d'analyse conduisent à un résultat cohérent lorsque cela est possible.

α "chanter IMP".

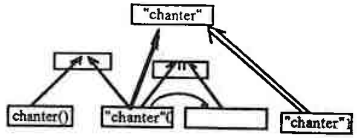
Supposons qu'il existe en mémoire une instance particulière de l'unité 'chanter' faisant partie d'une structure de cohérence (structurelle) où celle-ci est associée à un certain élément qui lui succède (cf figure). On peut voir cet élément comme étant une abstraction de différentes marques de conjugaisons lors de l'usage de 'chanter' dans différents contextes. Cette zone de cohérence fait alors elle-même partie d'une sur-zone représentant l'association générale d'une forme conjuguée de 'chanter' avec le prédicat correspondant (noté ici : chanter()). Nous avons donc la structure partielle:



De même, le suffixe 'IMP' est localement associé aux structures de la forme (P + IMP) au travers d'une de ses instances. Dans le cas présent cette structure va entrer dans une zone de cohérence similaire à la précédente, mais où le prédicat est inconnu, alors que certaines relations temporelles sont représentées (suivant la forme qui avait été donnée de l'imparfait). Nous avons donc une partie de la structure liée à 'IMP' :



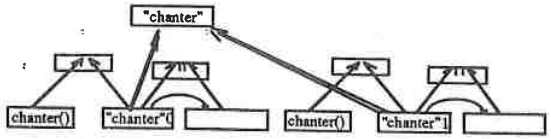
Le problème de la reconnaissance de la portion d'énoncé "chanter IMP" correspond à la mise en commun des informations contenues dans ces deux représentations. Le schéma d'analyse adopté est du type gauche-droite, c'est à dire dans l'ordre d'apparition (d'écoute) des différents éléments de l'énoncé. La reconnaissance du mot 'chanter' va permettre de relier l'occurrence particulière "chanter"1 à la classe générale "chanter":



La recherche de nouvelles propriétés pouvant être associées à "chanter"1 s'effectue le long de l'arborescence créée par la relation d'héritage. On aboutit ainsi au nœud général "chanter", représentant tous les homophones (ou les homographes quand l'énoncé est présenté sous forme de texte écrit). Deux solutions s'offrent alors, soit remonter l'arborescence pour rechercher un schéma plus général, soit redescendre le long d'une autre branche, pour voir si une autre instance (ou classe d'instance) de "chanter" ne présente pas une régularité qui pourrait être utilisée dans le cas présent.

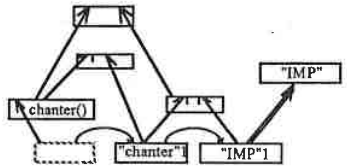
Nous avons vu (3.4.2) que les propriétés utiles étaient d'abord présentes au niveau des "cousins" le long de la hiérarchie. Pourtant, comme la condition d'appariement est seulement de posséder des ancêtres communs, il est possible d'aller explorer cette hiérarchie bien plus loin. Il est évident que le nombre de voies à explorer à un niveau supérieur sera très important avant d'aboutir à un résultat satisfaisant pour cette analyse. Nous n'allons donc détailler que l'exploration des branches utiles pour cet exemple.

Au niveau du nœud "chanter"0 existe donc une structure de cohérence qui, comme le nœud est un cousin de "chanter"1, peut être instanciée en prenant comme origines "chanter"0 et "chanter"1 (mécanisme présenté en 2). Le résultat de cette opération est alors une nouvelle structure liée de différentes façons aux éléments préexistants conformément au schéma (où certaines relations d'héritage ont été omises) :



A ce stade de l'analyse, l'instanciation d'une zone inconnue pose plusieurs problèmes de stratégie. En réalité deux solutions sont envisageables. Soit explorer les liens concernant cette zone inconnue de façon à préciser sa structure plus finement, soit, si cette information est disponible, chercher si l'entrée suivante peut convenir pour occuper cette place, ou servir de base à une sur-structure pouvant éventuellement occuper cette position.

Ces deux stratégies correspondent en fait au niveau local, à deux approches descendante ou ascendante. Dans le premier cas, on cherche d'autres informations contextuelles avant d'avancer dans l'analyse, alors que dans le second on choisit d'utiliser en priorité les portions d'énoncé disponibles le plus rapidement possible. En cas d'exploration en largeur, c'est la stratégie la plus rapide qui aboutira en premier. Dans le cas présent on trouvera "IMP" qui succède à "chanter". Par un mécanisme analogue à celui présenté au début de cet exemple, il est alors possible d'activer la zone de cohérence autour de "IMP"0, et de l'instancier en opérant une union avec la structure déjà existante. Le résultat de cet opération est la représentation suivante :

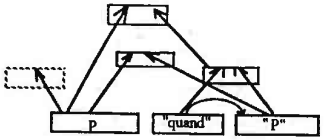


Nous pouvons remarquer que certaines contraintes non mentionnées doivent exister pour que cette dernière union soit possible. En particulier, nous avons vu que deux zones ne sont superposables que si elles possèdent un ancêtre commun le long de la hiérarchie. Ceci vaut pour les zones inconnues et les zones instanciées qui ont été mises en correspondance.

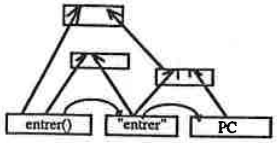
Une analyse équivalente aurait été possible en partant de l'élément "IMP" dans une perspective plus proche des problèmes de parole (analyse par îlot de confiance par exemple). Le résultat aurait bien évidemment été le même à la différence que les deux zones de cohérence utilisées dans cet exemple auraient été instanciées dans l'autre ordre, et une éventuelle prédiction aurait porté sur "chanter" en remontant la relation d'adjacence.

β "quand entrer PC".

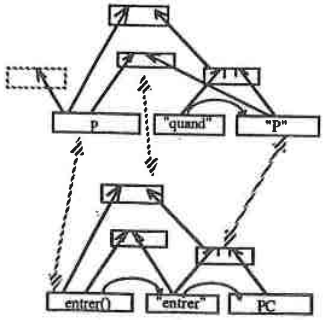
Nous traitons maintenant la deuxième partie de l'énoncé en portant cette fois plus notre attention sur le mécanisme d'appariement par rapport aux liens d'héritage que nous avions considérés précédemment. Nous pouvons tout d'abord proposer une structure possible autour de "quand" (une instance), où l'action qui suit ce mot est contenue dans une autre qui reste à déterminer (donc une zone d'attention). Par souci de clarté, nous avons choisi une structure de la forme "quand" + "P" et nous avons étiquetés les zones en conséquence :



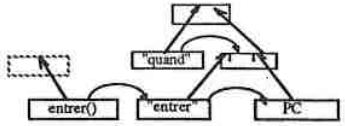
La prédiction de "quelque chose" derrière le mot "quand" entraîne la reconnaissance et l'analyse ; c'est ainsi que l'on obtient pour "entrer PC" une construction naturelle dont nous ne donnons pas plus les étapes :



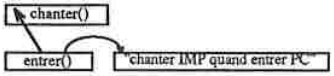
Une fois ces deux représentations construites, il reste à les apparier. Ce processus est facilement envisageable par l'intermédiaire des structures syntaxiques, de sorte que la zone de cohérence entre "entrer" et PC s'apparie avec la zone étiquetée par "p" dans le dessin précédent. Il est alors possible de visualiser cet appariement en pointillés :



La vérification de la cohérence ne posant pas de problème ici, nous pouvons donc donner une représentation de la proposition complète "quand entrer PC" (nous avons omis certaines zones de cohérence) :



Enfin, il ne reste plus qu'à construire la représentation complète de l'énoncé "chanter IMP quand entrer PC" grâce aux deux zones d'attention que l'on peut apparier respectivement avec 'entrer()' et 'chanter()' pour obtenir la structure finale suivante :



3.6 APPLICATIONS

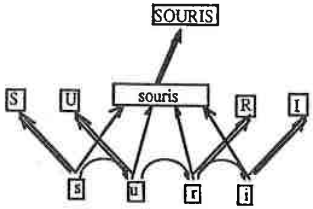
Nous présentons ici certains domaines d'application de notre modèle autour du dialogue oral homme-machine. Après avoir détaillé un modèle général, nous le situons par rapport à des activités que nous désirons unifier en montrant que les mécanismes qui interviennent sont finalement assez proches.

3.6.1 Temps et langage.

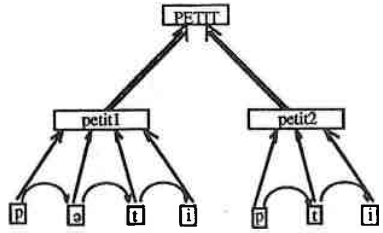
1 LEXIQUE ET PHONOLOGIE.

Dans notre modèle de représentation temporelle, les mots du lexique, comme toutes les autres unités du langage, sont représentés comme des zones particulières. Cependant, comme toutes les autres zones, elles sont décomposables en unités plus fines, dès que l'on désire connaître avec précision leur structure temporelle. C'est pourquoi l'analyse phonologique d'un mot peut être intégrée facilement dans le cadre de notre modèle.

Prenons un mot simple : "souris" ; son patron phonétique est /s u r i/, ce qui, sous forme temporelle peut être représenté ainsi :



Une instance particulière de ce nous pourrions appeler l'entrée lexicale 'SOURIS' est décomposée sous la forme de quatre instances des phonèmes 'S' 'U' 'R' et 'I'. Nous savons maintenant que les propriétés temporelles d'une classe de zones passent souvent par ses instances ; il n'est donc pas étonnant de ne rien connaître directement à propos de 'SOURIS'. De plus, il peut arriver que des variantes apparaissent pour un même mot qui peut se prononcer suivant les cas de différentes manières. Ainsi, le mot "petit" se représente de la manière suivante, où la deuxième instance de 'PETIT' correspond à une prononciation où le 'ə' a été élidé.



Les représentations ci-dessus sont utilisables dans deux sens différents suivant que l'on désire générer ou reconnaître la représentation phonétique d'un mot. Quand en entrée on dispose d'un phonème perçu, il faut définir un moyen de le rattacher dans un premier temps à sa classe d'origine. Ce mécanisme fait partie de ce que nous avions appelé des opérations *non-conscientes*, c'est-à-dire ne faisant pas intervenir de facultés de raisonnement. Nous supposerons que nous disposons d'un décodage acoustico-phonétique ou même d'un réseau de neurones qui fournit un résultat correct (pas nécessairement parfait) pour notre système.

Une fois un phonème disponible, son intégration normale au système tout entier va passer par un parcours de l'arbre d'héritage pour retrouver des structures dans lesquelles il pourra s'insérer. Cette phase va notamment conduire à des prédictions relatives aux autres éléments de ces structures. Ainsi si en entrée on a reconnu le phonème /s/, on voit que par l'intermédiaire de l'arborescence partielle qui a été proposée pour "souris", des prédictions phonétiques pourront être réalisées pour les phonèmes /u/, /r/ et /i/. Bien plus, quand des structures plus courantes sont mémorisées dans le système (des diphones par exemple), on se rapproche étonnement du modèle en cohorte de Marslen-Wilson (80).

Inversement, quand un mot apparaît dans une analyse, comme résultat d'une prédiction lexicale par exemple, il peut être décomposé de différentes manières au niveau phonétique et conduire à autant de vérifications sur le signal. Là encore, des factorisations éventuelles, qui sont facilement réalisables dans notre modèle, peuvent toujours simplifier cette opération ({/t/ /i/} dans "petit" et "p'tit" par exemple).

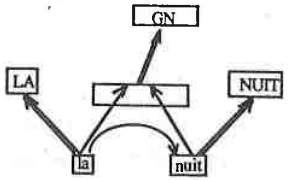
Nous avons pris l'option de choisir les phonèmes comme éléments minimaux d'un système de dialogue oral homme-machine ; il est évident que ce choix n'est pas le seul possible et peut varier suivant le niveau de décodage dont on dispose ou la langue particulière qui sert pour les échanges. Ainsi, on peut choisir le mot comme unité

minimale dans le cas d'une lecture de texte (il y a peu d'ambiguïtés possibles) ou le diphone si la langue utilisée est le Japonais par exemple, où le langage est structuré en syllabes (a, ki, su, te, mo, etc...) de façon très rigoureuse.

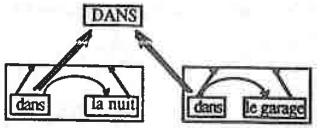
Enfin, la prosodie peut se révéler importante pour déterminer a priori les ensembles d'informations qui constitueront des zones de cohérence comme dans "Nancy bat Lens (*balance*) 2 buts à 1", la détermination de frontières adéquates peut éviter une recherche sémantique un peu plus longue.

2 SYNTAXE.

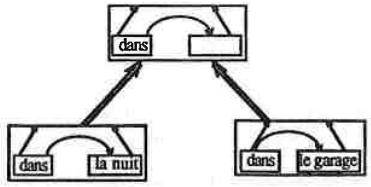
La syntaxe, au sens le plus stricte du terme, concerne les associations des mots dans une phrase de façon purement structurelle. En ce sens, sa représentation dans notre modèle se rapproche beaucoup du traitement de la phonologie que nous avons déjà proposé. En effet, il s'agit de représenter comment certains mots (donc des zones temporelles) peuvent être associés au sein d'une même zone de cohérence. Ainsi, le groupement "la nuit" peut être représenté sous la forme d'une instance particulière d'une zone générale qui serait, à la rigueur, un "groupe nominal" :



Si nous voulons maintenant exprimer que ce groupement, ainsi que "le garage" peuvent s'insérer dans une structure plus complète en étant précédés par "dans", nous pouvons arriver à la situation suivante :



Après une phase d'apprentissage, les deux zones de cohérence peuvent être réunies sous un même ancêtre représentant les structures de la forme "dans" + "..." :



Nous pouvons maintenant dire quelques mots au sujet des problèmes d'accord entre les différents éléments d'une phrase. Dans notre modèle, il n'existe pas de catégories syntaxiques. Celles-ci sont traduites indirectement par des classes de mots qui interviennent plus régulièrement dans des structures particulières. Ce choix est volontaire, puisqu'il permet à un système reposant sur notre modèle de réagir face à des nouvelles structures syntaxiques. Le problème des accords se pose donc par rapport à ces classes de mots qui ont été éventuellement apprises au cours des expériences antérieures du système. Cependant, une phase d'initialisation est nécessaire et c'est à ce moment que des catégories seront introduites (de force) dans le système. Prenons l'exemple des groupes "la nuit" et "le garage". A un certain niveau d'abstraction, il s'agit de la même structure (Déterminant + Nom), mais si on s'attache aux instances de cette structure, on peut en distinguer deux classes principales : les formes (Déterminant Masc. + Nom Masc.) et (Déterminant Fém. + Nom Fém.), chacune de ces sous-catégories syntaxiques étant des instances de 'Déterminant' et 'Nom' respectivement. La figure 3.4 présente un bilan de ces problèmes sous la forme d'une unique représentation. Certains nœuds de l'arborescence classe-instance ont été étiquetés avec des catégories syntaxiques fictives correspondant à ce que peut avoir en tête un concepteur de système mettant au point la base de connaissance initiale.

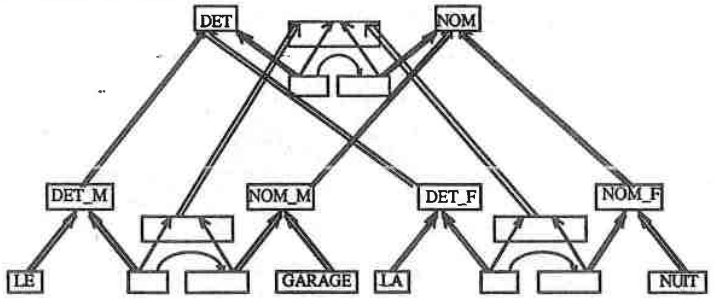


Figure 3.4 : représentation de certains groupes nominaux simples.

3.6.2 Le temps signifié.

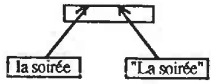
1 LES EXPRESSIONS TEMPORELLES.

De nombreuses parties du discours expriment des valeurs temporelles qui peuvent être représentées dans notre modèle par des zones. Nous n'allons pas faire ici une étude linguistique exhaustive qui n'est pas de notre ressort, mais nous allons plutôt insister sur quelques problèmes qui apparaissent.

En dehors des formes prédicatives contenues dans des propositions, noms ou adjectifs, une grande classe d'expressions linguistiques est constituée de groupes nominaux et d'adverbes faisant référence directement à des intervalles de temps (au sens large) que nous manipulons tous les jours. Ainsi, "lundi", "la soirée" ou "hier" sont directement compréhensibles en tant que zones temporelles.

Durant l'étude de l'adverbe "hier", nous avons déjà observé le flou qui entoure sa description temporelle exacte, à la fois pour ce qui est de sa position précise dans le temps ou de sa durée effective, en fonction du degré de connaissance de l'auditeur. En effet, la durée, tout comme la position, est définie dans notre modèle de manière exclusivement relative et dépend des schémas qui peuvent être instanciés autour d'une zone donnée. Finalement, le langage, seul, n'exprime que peu de choses au niveau temporel et il ne donne souvent que des clés d'accès à des éléments de représentations qu'il est toujours nécessaire de compléter.

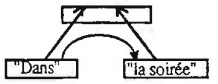
Nous ne traiterons ici que l'exemple de l'expression "dans la soirée" pour montrer l'influence du contexte d'analyse sur sa possible valeur temporelle. Toute seule, "la soirée" n'exprime qu'une zone temporelle grâce à l'intervention désormais classique d'une zone de cohérence :



Le mot "dans" s'insère quant à lui dans une expression plus complexe au niveau syntaxique et, si on ne garde que sa valeur temporelle, il situe (au niveau sémantique) un événement à l'intérieur de la zone désignée par ce qui le suit. Cette action est inconnue,

tant que d'autres expressions (cf "il avait neigé") ne viennent pas l'expliciter ; c'est donc une zone d'attention.

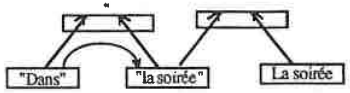
La construction d'une représentation mettant en œuvre toutes les informations précédentes relatives à "dans" se fait par étapes, en fonction des instanciations successives. Nous choisissons maintenant une séquence possible pour reconnaître "Dans la soirée". En premier lieu, l'apparition de "dans" va générer un schéma structurel (pour ne pas dire "syntaxique") dans lequel il pourra s'insérer, ainsi que : "la soirée" au moment où cette dernière expression est reconnue (on ne précise pas plus l'analyse ici). Localement, le schéma obtenu est donc :



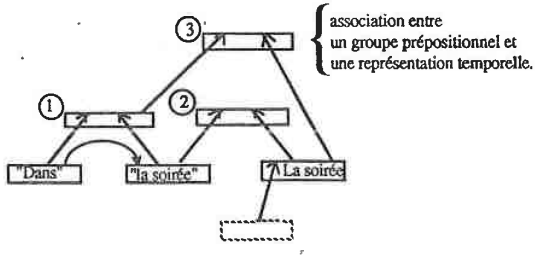
cohérence structurelle

éléments lingistiques reconnus

A ce stade, on peut choisir de détailler la zone de cohérence que l'on vient de construire ou d'essayer d'insérer "la soirée" dans un schéma de cohérence pour en trouver la signification. Dans ce dernier cas, on obtient :

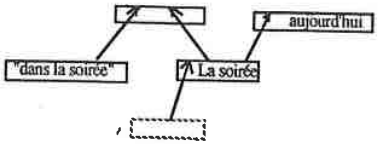


Si maintenant on met en œuvre un schéma associant à une structure syntaxique de la forme "dans P" une représentation où l'on situe un événement à l'intérieur de la zone associée à "P", la représentation finale de "Dans la soirée" devient :

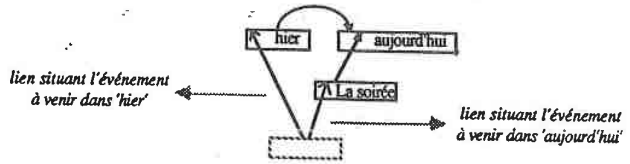


Dans le schéma précédent sont numérotées les trois "cohérences" qui sont intervenues successivement (1, 2 puis 3) dans notre analyse. Cet ordre a permis notamment à zone '3' de s'insérer naturellement au dessus des zones déjà existantes. Un autre processus d'analyse, à savoir [1, 3, 2], correspondant à l'explicitation de la structure syntaxique induite avant la recherche de la signification de ses composants, aurait conduit à *prédire* par avance une zone associée au message "la soirée". Le schéma '2' se serait alors naturellement apparié avec cette prédiction.

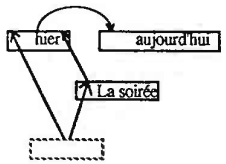
L'analyse purement "linguistique" fournie ci-dessus ne donne cependant aucune indication sur la position temporelle de la zone associée à "la soirée" (que nous notons 'la soirée'). Pour obtenir plus de renseignements, il faut encore mettre en œuvre d'autres schémas connus de l'auditeur. Ainsi, on peut imaginer qu'une soirée est toujours contenue dans un jour particulier et si le dernier jour dont on ait parlé est le jour courant, la représentation de "Dans la soirée" est immédiate :



Le contexte immédiat du discours peut éventuellement fournir des indications pour situer "la soirée". Ainsi, dans l'expression "Hier, dans la soirée", on observe deux localisations d'un événement à venir, ce qui réduit les possibilités d'interprétation de "la soirée". Par exemple, si nous instancions le même schéma que précédemment, nous obtenons une *incohérence* au niveau temporel ; comme cela apparaît dans la figure ci-dessous (la zone d'attention est à la fois située *dans* et *après* 'hier') :



La seule interprétation qui sera acceptée par le vérificateur de cohérence temporelle, après appariement, est donc :

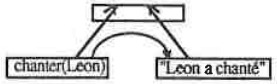


Nous observons là encore que peu d'informations dans ce schéma sont strictement *linguistiques*. Dans notre modèle, qui s'appuie beaucoup sur des schémas existants à n'importe quel niveau de connaissance, la langue est strictement indissociable de l'information qu'elle exprime.

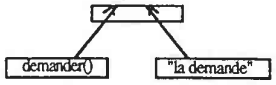
2 A PROPOS DES PRÉDICATS.

α Rappel.

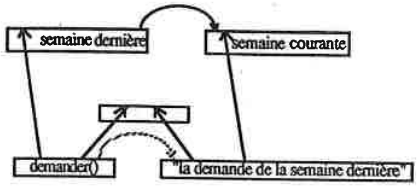
L'expression prédicative est une des façons privilégiées de traduire un événement en un acte linguistique. Inversement, toutes ces expressions sont associées, au niveau le plus élémentaire d'analyse, à des zones temporelles par un simple lien de cohérence. Si par hasard, des informations supplémentaires sont disponibles, comme dans : "Leon a chanté", il est même possible de situer la zone perceptive et la zone conceptuelle l'une par rapport à l'autre (schéma classique désormais) :



Seulement, il n'y a pas que les propositions simples qui génèrent des zones temporelles ; par exemple, un prédicat exprimé sous la forme d'un nom est aussi associé à un événement qu'il représente. Ainsi, à l'expression "la demande" peut être associée une zone temporelle de façon identique à ce qui avait été fait pour une expression comme "la soirée" :



Cette représentation est typiquement intemporelle, puisqu'aucun marqueur morphologique ne vient situer l'action par rapport au temps de l'énoncé[1]. Cependant, il peut exister des groupements secondaires qui explicitent la position de l'action dans le temps comme dans "la demande de la semaine dernière", qui peut se représenter ainsi :



Le lien d'adjacence en grisé dans cette représentation indique une relation calculée en fonction des autres contraintes du réseau de zones présenté. Cette relation n'est pas vraiment "dans les mots". Elle ne peut être connue qu'après avoir effectivement construit la structure temporelle portée par l'énoncé.

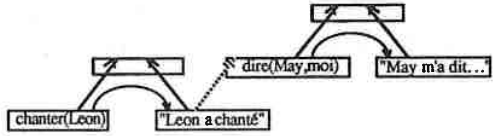
β Temps de conjugaison.

Le principe de base de notre analyse des temps de conjugaison est qu'à un marqueur linguistique simple ("ait" pour l'imparfait) ou complexe (passé composé) on peut associer certaines relations temporelles entre l'énoncé et la situation qu'il exprime, avec éventuellement une création de zones d'attention intermédiaires. Dans le formalisme de Hornstein, ces zones peuvent parfois être considérées comme des points de référence autour desquels se construit la représentation. Pourtant, on a vu l'origine profonde des zones d'attention qui sont des prédictions au sujet desquelles on ne possède pas d'informations. Dans le cas du point de référence 'R', on possède moins de justification de son existence, si ce n'est par l'intermédiaire de règles de construction qui lui sont associées.

Nous n'allons traiter ici que quelques points qui nous paraissent remarquables, en sachant bien que des études plus précises sont à faire relativement aux temps grammaticaux.

[1] On remarquera qu'en japonais certaines catégories d'adjectifs se "conjuguent" au passé, au même titre que les verbes et donc leur situation temporelle est mieux définie.

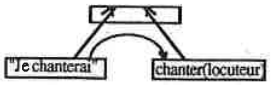
Pour l'instant, nous avons surtout porté notre attention sur les temps du passé qui, au niveau de la représentation, sont les plus simples à concevoir. Quand les zones de cohérence ont été introduites, nous avons aussi dit quelques mots du présent français dont certaines de ses multiples valeurs pouvaient être explicitées grâce à elles. Parmi ces valeurs le présent rapporté peut aussi être explicité sous la forme de zones temporelles, moyennant un certain flou sur la notion d'univers de croyance. Sans valeur de généralité, la phrase "May m'a dit : «Leon a chanté»" se représente approximativement ainsi :



On constate alors que la zone étiquetée 'dire(May,moi)' délimite un espace de référence à partir duquel il est possible de situer l'énoncé rapporté dans le temps. On remarque aussi que "Leon a chanté" est en quelque sorte une description plus fine de l'action 'dire(May,moi)', d'où la relation d'inclusion qui a été introduite entre les deux. Enfin, nous avons supposé dans tous les exemples traités jusqu'ici, que le locuteur était honnête[1], en ce sens que les informations qu'il apporte sont directement intégrables dans la base du système. Rejeter ce principe est impossible pour nous en l'état actuel de notre modèle, puisque nous n'avons pas développé la notion de possible, c'est-à-dire d'informations contradictoires (ou négatives) présentes en même temps dans la base de connaissance. Pour traiter ce problème, il serait nécessaire d'établir pour chaque locuteur et à la limite pour chaque énoncé de celui-ci un monde possible (non pas au sens de la logique modale, mais au sens de Eco ou Fauconnier) dans lequel serait représentée l'information. Dans notre modèle, la définition de ces mondes passe vraisemblablement par un mécanisme similaire aux zones de cohérence, mais que nous n'avons pas modélisé pour l'instant. Dans l'exemple précédent, on aurait ainsi deux espaces de ce type qui s'emboiteraient pour les deux niveaux d'énoncé qui apparaissent.

Le même problème existe pour la compréhension d'un temps dont nous n'avons pas encore parlé : le *futur*. En première approximation et en supposant que le locuteur soit infaillible, la phrase "je chanterai" peut se représenter ainsi :

[1] On peut à juste titre considérer cette hypothèse comme valide dans le cas d'un dialogue homme-machine finalisé : l'objectif du locuteur est en fait d'obtenir satisfaction pour sa demande et non de mettre en défaut le système.



Pourtant, nous n'avons pas de moyen de gérer l'existence de plusieurs futurs possibles par cette représentation qui inclut directement l'information 'chanter(locuteur)' dans la base de connaissances. De façon générale, nous pouvons remarquer que c'est encore le même type de difficulté que précédemment et il semble bien que la solution passe par une seule méthode de traitement. Nous n'allons pas tarder à voir que d'autres niveaux de connaissance sont aussi touchés par ce phénomène.

γ Actes, processus et autres objets divers.

Dans notre discussion sur les prédicats, nous n'avons pas abordé des notions qui sont couramment rencontrées (particulièrement chez les linguistes) et qui concernent la valeur temporelle précise d'un événement particulier auquel on fait référence dans un énoncé. Ainsi, on rencontre souvent les termes de 'acte', 'processus' ou 'état' (et parfois d'autres) pour désigner différentes catégories qui, semble-t-il, se distinguent parmi toutes les situations possibles. Cette classification marque en fait deux phénomènes distincts. Souvent, cela correspond à une manière d'envisager ces événements alors que la réalité sous-jacente est plus complexe. L'exemple que nous avions donné au 3.2.2 ("hier, j'ai marché toute la journée") est significatif puisque suivant les cas, on distinguera plutôt l'acte de marcher ("je suis allé de X à Y"), le processus ("Plus je marchais, plus j'avais soif") ou même l'état ("J'étais tranquillement en train de marcher, quand une gazelle est apparue"). C'est donc plus une notion d'aspect qui est présentée dans cette description et donc des problèmes qui touchent le langage particulièrement (par exemple la différence passé-composé/imparfait).

A l'opposé, des ensembles de situations possèdent des caractéristiques temporelles très semblables qui font penser qu'elles doivent être réunies en catégories stables. Il ne s'agit plus là de problèmes linguistiques, mais surtout de la manière dont sont perçues au niveau cognitif ces situations, indépendamment des mots utilisés pour les exprimer. Pourtant la classification correspondante est trop souvent imposée à la langue comme nous l'avons déjà constaté dans les grammaires de cas par exemple. De plus, la difficulté de situer avec précision certains verbes à la limite entre deux catégories, prouve bien que celles-ci (qu'il y en ait deux, trois, quatre ou plus) ne sont qu'une représentation réduite de phénomènes

plus complexes et, comme beaucoup de représentations figées, ce découpage interdit qu'un système considère un événement de manière plus fine que celle que cet arbitraire autorise.

Notre modèle se prête beaucoup plus à une description aussi fine que nécessaire des événements de l'univers du discours, qu'ils soient exprimés ou non. A un niveau important d'abstraction et seulement si ces catégories se rencontrent réellement, on pourra rencontrer des zones temporelles dont la structure (les propriétés temporelles) correspond à un acte, un processus ou un état. A la limite, ceci pourra résulter d'une initialisation de la base de connaissances à partir d'une étude linguistique qui aura été faite. Cependant, les différentes instances d'événements qui vont apparaître dans l'expérience du système feront qu'il existera toujours des sous-catégories de plus en plus fines qui donneront à celui-ci une très grande souplesse de description des phénomènes. Notre modèle présente donc cet avantage de ne pas dépendre obligatoirement de catégories immuables, bien que celles-ci puissent toujours servir de base pour une première analyse.

3.6.3 Raisonnement dans l'univers du discours.

1' DIALOGUE ET STRUCTURES DISCURSIVES.

Les structures de dialogue telles qu'elles ont été considérées dans le projet initial présenté au premier chapitre ([Roussanaly 88]) peuvent être resituées dans l'optique de notre modèle. L'analyse fonctionnelle du dialogue repose, au sens de Moeschler sur les notions d'actes, d'interventions et d'échanges qui concernent tout aussi bien le système que le locuteur dans un dialogue homme-machine. Si on choisit de considérer le traitement opéré par la machine, il faut retrouver pour chaque intervention du locuteur le type d'échange auquel elle correspond et suivant les cas répondre en conséquence. Pour faciliter cette tâche, A. Roussanaly (88) a proposé une typologie des actes de langage spécifique à l'application de renseignements administratifs qui introduit par exemple les notions d'ouverture/clôture de dialogue, production d'informations, contestation, question, etc...

Par ailleurs, l'analyse des échanges montre combien ceux-ci sont enchâssés de façon récursive comme le montre la figure 3.5.

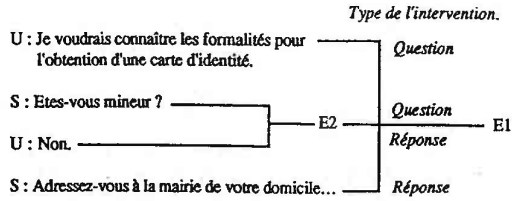


Figure 3.5 : Structure d'échanges (d'après [Roussanaly 88]).

A partir de ces données, on perçoit immédiatement comment les interventions et les échanges peuvent être représentés au niveau temporel par des zones. Nous n'allons pas ici décrire en détail un exemple, mais commenter les deux aspects principaux de cette approche :

- chaque énoncé d'un dialogue est assimilable dans notre modèle à un ensemble cohérent d'informations et donc à une zone temporelle. Dès lors, la typologie des actes de dialogue peut être interprétée comme la définition de certaines classes de zones dont la pertinence est nette dans un dialogue finalisé. Associer un énoncé à une classe consiste alors à retrouver dans une structure plus fine de celui-ci des propriétés temporelles (structure syntaxique d'une question par exemple) allouées par avance à cette classe. La discussion entamée pour les catégories d'actes, de processus et d'états peut être réitérée ici, puisque la typologie des actes de dialogue n'a pas dans notre modèle de caractère immuable, mais peut être remodelée au fur et à mesure que le système *apprend* de nouvelles structures d'énoncés.

- les échanges, à leur tour peuvent être assimilés à des zones de plus haut niveau encore, décomposables sur la bases des interventions. On retrouvera ainsi sûrement des associations standards du type question/réponse, mais d'autres plus complexes peuvent être enseignées à un système ou apprises par celui-ci.

Reste un problème auquel nous ne donnons pas de réponse actuellement qui est le passage à l'acte linguistique pour le système. Dans le cas de perceptions, nous avions considéré que des éléments successifs sous forme de zones temporelles arrivaient au système pour être ensuite analysés. Inversement, il faudrait considérer un mécanisme qui déclenche l'émission d'un énoncé par la machine dès que certaines conditions sur les structures de zones sont rencontrées. C'est là encore un axe de recherche à prendre en compte dans l'avenir.

2 PLANIFICATION.

Le problème de la planification dans un dialogue homme-machine consiste simplement à raisonner au plus haut niveau sur les événements de l'univers du discours. Un événement étant représenté par une zone temporelle, toute structure de raisonnement (cause-conséquence par exemple) s'assimile à une zone de cohérence sur ces événements. Il ne s'agit pas alors pour nous de définir des structures figées telles que les scripts de Schank (77), mais par exemple des schémas de raisonnement plus réduits pouvant servir de base à la construction de plans plus complets par instanciations successives.

De même une planification trop figée telle qu'elle est envisagée dans STRIPS ([Fikes 71], voir aussi [Nilsson 80]) ne correspond pas à une réalité plus complexe d'opérateurs et d'états qui se déroulent dans le temps de manière pas nécessairement séquentielle. Bien qu'une possibilité de hiérarchie ait été envisagée dans ABSTRIPS ([Sacerdoti 74]), elle correspond plus à un traitement mécanique des opérateurs pris en compte, par comparaison avec la hiérarchie introduite par l'apprentissage de zones de cohérence, tel que nous l'avons proposé.

Parmi les domaines d'application de notre modèle, la planification semble privilégiée car elle fait intervenir toutes les propriétés temporelles et d'héritage que nous avons étudiées. De plus, des prévisions sur un univers particulier s'opposent souvent dans le temps, ce qui impose là encore la prise en considération des possibles pour gérer ces problèmes convenablement. Les quelques développements précédents ont simplement situé la manière dont nous envisagions le problème, sans vraiment donner de solutions. Une analyse complémentaire reste à faire sur ce sujet.

3.6.4 La mémoire et l'oubli.

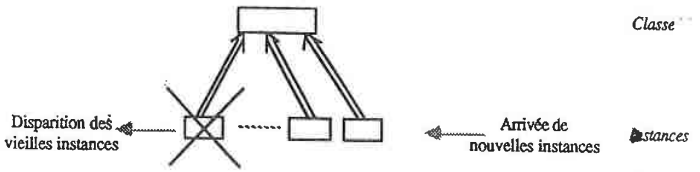
Dans notre modèle de représentation des informations temporelles, nous avons prévu des mécanismes pour constamment fabriquer de nouvelles zones, soit par apprentissage, soit par instanciation de zones pour compléter les propriétés d'une zone particulière. Il se pose alors le problème de l'encombrement de la mémoire d'un système fonctionnant sur ce modèle et donc celui de l'oubli. Pour l'homme aussi, l'oubli est nécessaire dans son fonctionnement cognitif continuel. A titre d'illustration, nous pouvons rappeler le cas d'un sujet mnémoniste cité par S. Rose (75) qui "souffrait de pouvoir se souvenir de tout ; tout ce qu'il voyait, il l'emmagasinait. Parler avec lui provoquait dans son esprit des

trains continus de correspondances" (p.281). Contrairement à ce que l'on pourrait penser une trop grande mémoire n'est pas forcément un avantage, ainsi, "il ne pouvait assumer de fonctions de façon adéquate, parce qu'il trouvait difficile de comprendre ce qui lui était dit, à cause des rappels en chaîne que même une phrase très simple suscitait. En désespoir de cause, il devint un homme-mémoire professionnel au théâtre".

Par analogie, un système automatique qui emmagasinerait trop d'informations serait vite débordé par les associations possibles que l'apparition d'un simple mot pourrait générer. Nous avons d'ailleurs déjà rencontré ce problème quand nous parlions des stratégies d'accès aux "cousins" d'une instance donnée. Plus ces "cousins" sont nombreux, plus les problèmes de stratégie risquent de devenir cruciaux pour un système. Une solution à ce problème est alors d'envisager un mécanisme d'oubli *modéré* qui "nettoierait" régulièrement les zones inutiles.

Un tel mécanisme doit prendre en compte deux phénomènes complémentaires : d'un côté les zones les plus anciennes doivent disparaître pour laisser la place aux plus récentes, d'un autre les propriétés attachées à ces zones doivent être conservées dès qu'elles sont reconnues régulièrement par le système. Dans le cas de notre modèle, le mécanisme d'apprentissage développé permet de restreindre ces conditions à la première qui autorise d'elle-même un allègement cohérent de la mémoire.

Le mécanisme de base consiste donc à garder une trace, pour une classe de zones donnée, du moment d'arrivée de chacune de ses instances[1]. Un processus automatique élimine alors les instances trop vieilles suivant le schéma suivant :



Sur la base de ce schéma général, nous observons que les deux opérations d'apprentissage et d'instanciation (donc de remémoration) ralentissent le phénomène

[1] Le temps dont il est question ici n'a pas de rapport avec le temps de notre modèle, ce serait plutôt un temps "biologique" qui n'est pas représentable par les zones temporelles (il y aurait là un problème d'auto-référence).

d'oubli pour les propriétés temporelles qui sont ainsi manipulées. L'importance du rappel est un phénomène qui possède une réalité psychologique puisque, comme le signale S. Rose (75 : p.277) : "Le rappel n'efface pas le souvenir de la réserve ; au contraire, plus on se rappelle une chose, mieux elle semble être fixée". Pour notre modèle cela présente l'avantage de pouvoir réutiliser longtemps des informations usuelles, dont particulièrement : le langage. Observons comment cela se passe :

- lors d'un apprentissage qui met en œuvre deux zones temporelles qui peuvent être anciennes (attention aux deux notions de temps!), la *nouvelle* zone créée devient une instance très récente des ancêtres communs à toutes deux. De la sorte, leurs propriétés communes acquièrent une nouvelle fraîcheur qui les perpétue ainsi pour les raisonnements futurs du système.

- pareillement, une instanciation associe des propriétés temporelles plus anciennes à une nouvelle zone. Celles-ci sont donc là-encore renforcées.

Finalement, alors qu'un mécanisme d'oubli semble fondamental pour un système qui apprend et donc qui construit beaucoup d'objets, il s'avère qu'un tel mécanisme est facilement pris en compte dans notre système. Une difficulté éventuelle réside dans la gestion de la vitesse d'oubli qui dépend étroitement de la capacité de reconstruction du système.

3.7 Conclusions.

Dans ce chapitre, nous avons exposé les principes de base d'un modèle cognitif de représentation des informations temporelles pouvant apparaître dans un dialogue homme-machine. Ce modèle repose en fait sur deux aspects complémentaires qu'il est important de percevoir pour juger des développements futurs auxquels il est susceptible de conduire.

En premier lieu, nous avons défini un ensemble d'objets temporels fondamentaux sur lesquels peut s'appuyer notre modèle. Ce sont bien évidemment les zones temporelles, mais aussi les relations qui les lient, à savoir l'adjacence, l'inclusion et la relation d'héritage. Pour ces objets des opérations spécifiques ont été fournies permettant notamment de vérifier la cohérence temporelle ou la construction d'objets nouveaux le long de la hiérarchie des zones. Cette partie de notre modèle, comme outil fondamental, est relativement figé et ne doit pas subir de grands bouleversements à moyen terme.

A l'opposé, pour les différents domaines d'application de notre modèle, nous avons proposé des structures de représentation de mots, syntagmes ou phrases, qui sont beaucoup plus sujettes à critique, car elles résultent de la seule évaluation de l'auteur de ce texte, dont la culture linguistique ou psychologique n'est pas forcément suffisante. C'est pourquoi les développements ultérieurs donneront sûrement lieu à de nouvelles représentations pour les mêmes éléments linguistiques sans que cela remette en cause notre modèle. Au contraire, comme nous l'avons déjà laissé entendre, une étude approfondie serait nécessaire avec l'aide de linguistes par exemple pour essayer de valider les outils utilisés. En effet, contrairement à un modèle logique où l'on cherche à montrer des notions telles que la cohérence ou la validité, un modèle comme le nôtre ne sera "prouvé" que par un confrontation avec des données *réelles* qu'il pourra, ou non, prendre en compte.

De ce fait, la quantité de travail restant à fournir sur le seul domaine des informations temporelles dépasse de beaucoup ce que nous avons pu déjà faire. Le modèle proposé n'est donc pas un ensemble complet et fermé, mais un espace ouvert pour des recherches à venir.

4 Outils de manipulation de Zones Temporelles.

4.1 Introduction.

Nous décrivons dans ce chapitre l'état actuel de l'implémentation de notre modèle de représentation du temps. Nous avions insisté au troisième chapitre sur l'existence de deux niveaux de description du modèle :

- le support du modèle, constitué des zones temporelles et des relations d'adjacence, d'inclusion et d'héritage. Sur ces objets nous avons défini des mécanismes de base de vérification de cohérence et d'instanciation/abstraction de zones qui donnent au modèle ses possibilités de représentation et de déduction.

- le contenu d'un système reposant sur ce modèle, sous la forme de schémas standards de la langue et des événements de l'univers (par l'intermédiaire des zones de cohérence). A ce niveau, nous ajoutons les problèmes de stratégie dont les détails doivent encore être précisés.

Il est important de mettre en place dans un premier temps des outils ouverts de manipulation de zones temporelles, sur lesquels seront développées diverses configurations de travail dans le cadre éventuel d'applications plus ciblées. Il n'est donc pas question que nous implémentions directement notre modèle pour un domaine spécifique, mais nous voulons plutôt définir un espace de travail où un maximum de possibilités sont offertes à l'utilisateur.

Jusqu'à présent, nous avons essentiellement présenté les zones temporelles sous forme graphique, car ainsi il était beaucoup plus facile de les situer les unes par rapport aux autres et de suivre les mécanismes d'appariement qui ont été décrits. Dans le cadre d'une implémentation, le même problème se pose, puisqu'il faut comprendre à tout moment ce que fait le système et éventuellement lui apporter des connaissances qui s'intègrent bien avec les données déjà existantes. C'est pourquoi nous avons concentré une grande partie de nos efforts sur la mise en place d'une interface graphique complète, avant même d'implanter les opérations élémentaires sur les zones.

L'outil en cours de développement dont nous décrivons ici la structure est avant tout un *atelier graphique de manipulation de zones temporelles* dont les fonctionnalités évolueront en parallèle avec les développements ultérieurs (plus théoriques) du modèle. Dans ce chapitre, nous exposerons successivement la structure générale de cet atelier, puis les grandes fonctionnalités déjà disponibles, sachant que celles-ci sont à l'heure actuelle en constante évolution. Enfin, nous présenterons certains objectifs, concernant notamment les stratégies de recherche le long des relations d'héritage.

4.2 Environnement de développement.

4.2.1 Langage utilisé.

Comme pour le gestionnaire d'hypothèses lexicales présenté au premier chapitre, nous développons les outils de manipulation de zones temporelles en Flavors, sur-langage orienté objet de Lisp. Cependant, par souci de compatibilité et d'efficacité, nous utilisons la version accompagnant le CommonLisp fourni pour les stations de travail Sun. De plus, l'environnement choisi possède les primitives graphiques indispensables pour la mise en œuvre d'un système interactif.

Nous avons déjà observé comment la programmation en langage orienté objets permettait de créer des applications particulièrement modulaires et faciles à maintenir. Dans le cas de la définition d'un outil ouvert, c'est un aspect d'autant plus avantageux puisque nous désirons pouvoir modifier continuellement les possibilités de notre programme sans pour autant bouleverser sa structure d'ensemble. Nous allons maintenant voir comment le lien entre notre modèle et l'environnement de Flavors a été réalisé.

4.2.2 Représentation des zones temporelles sous Flavors.

1 Quelques problemes.

Lors de l'implémentation du gestionnaire d'hypothèses lexicales, nous avions assimilé le lien d'héritage entre deux nœuds d'une arborescence à la relation explicite qui existe entre les flavors dans le langage support. La même question se pose ici, mais nous devons y répondre autrement de par la complexité accrue des phénomènes à représenter. Dans le cas du gestionnaire, la relation d'héritage servait essentiellement à structurer les valeurs contenues dans le lexique et à l'occasion, à propager certaines informations ou des fonctions de nœud en nœud, sans que la nature profonde du lien d'héritage ait une

quelconque importance. Par contre, la relation introduite entre deux zones temporelles possède ses propres mécanismes de fonctionnement et demande donc à être manipulée explicitement, en particulier lors des opérations de création de nouvelles classes. De fait, contrairement à ce qui se passait pour le gestionnaire d'hypothèses où la structure du lexique était définie une fois pour toute, la hiérarchie des zones temporelles est dynamique. Par conséquent, la relation d'héritage sera implantée séparément de la structure du langage, ce qui assure d'ailleurs à l'ensemble une meilleure portabilité.

Un deuxième problème à résoudre concerne la représentation des relations en général comme des objets explicites, plutôt que de les sous-entendre, en indiquant pour chaque zone temporelle quelles zones lui sont reliées. Or nous avons vu que pour certaines opérations sur ces relations (ch.3.5), il était fait référence aux arcs entre deux zones plutôt qu'aux zones elles-mêmes. Bien plus, dans le cas de l'unification de deux zones, les deux modes de représentation sont nécessaires. Nous avons donc choisi, malgré la redondance qui en résulte, de traduire les relations, ainsi que les zones temporelles sous la forme d'objets manipulés par le système.

2 Structure de l'environnement.

Dans l'environnement choisi, tout objet intervenant dans la structure du système est représenté sous la forme d'une instance de flavor. Les différents flavors qui sont alors introduits représentent les classes d'objets de même nature. Par ce biais, nous obtenons une structure homogène dont la hiérarchie (au sens des flavors) est représentée dans la figure 4.1. Pour chaque flavor sont indiquées les variables d'instances qui lui sont attachées. Dans cette architecture, tous les flavors héritent d'une même classe, le flavor 'f_système' qui lui-même ne dépend que du flavor '*vanilla*' (flavor fondamental du langage). Au flavor 'f_système' est attachée la variable 'texte' qui sera ainsi une variable d'instance de tous les flavors du système. De façon générale, cette variable contiendra un commentaire ou une étiquette permettant de reconnaître la zone à l'affichage. Cependant, pour les zones de bas niveau ('*f_texte*'), elle représentera le texte effectif attaché à une d'entre elles. Il s'agira par exemple du graphème d'un mot pour que celui-ci puisse être reconnu et relié à la bonne classe de zones temporelles.

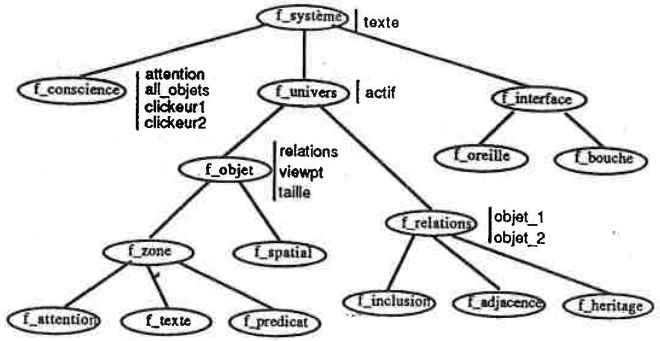


Figure 4.1 : une vision d'ensemble du système.

Au niveau le plus haut de l'arborescence on distingue trois catégories principales de Flavors :

- le flavor *'f_conscience'* délimite la partie active du système. C'est une instance de ce flavor qui va traiter les différentes commandes et déterminer les zones qui seront manipulées au fur et à mesure de l'avancement du raisonnement. A ce flavor sont attachées 4 variables principales : la variable 'attention' conserve à tout instant l'ensemble des zones candidates pour un traitement, et correspond approximativement à la mémoire à court terme du système. Pour l'instant elle est gérée comme une liste dont le premier élément est la zone la plus récente et donc la plus active. La variable 'all_objets' est une variable d'aide au développement contenant tous les objets du système ; nous ne pensons pas la conserver dans une version future. Enfin, les deux variables 'clickeur1' et 'clickeur2' contiennent les zones éventuelles qui ont été désignées par la souris dans l'interface graphique. Ces deux objets vont être appariés, ainsi que les zones qui les entourent. Suivant les cas, il y aura instanciation ou abstraction du résultat obtenu.

- le flavor *'f_univers'* est la racine de toutes les entités introduites par notre modèle, à savoir les objets de l'univers (*'f_objet'*) et toutes les relations sur ceux-ci (*'f_relations'*). La variable actif correspondante signale si l'objet en question doit être tracé ou non, ce qui correspond souvent à la présence de celui-ci dans la conscience.

- le flavor *'f_interface'* réunit tous les éléments du système chargés d'échanger des informations non-graphiques avec l'extérieur. Il contient une entrée (*'f_oreille'*)et une sortie (*'f_bouche'*) qui ne sont que des flux de caractères, puisque nous ne travaillons pour l'instant que sur les mots comme éléments les plus élémentaires de notre modèle (éléments perceptifs).

Les objets de l'univers contiennent trois informations qui déterminent leur comportement à la fois conceptuel et graphique. La variable 'relations' contient l'ensemble des relations d'inclusion, d'adjacence ou d'héritage que l'objet établit avec d'autres, sous la forme d'une simple liste d'instances d'une des trois classes de relations. Les deux autres variables concernent la longueur et la largeur du rectangle maximal affiché pour cet objet ('taille'), ainsi que le *viewport* attaché à lui, c'est à dire une zone graphique utilisée par l'environnement *Sun* pour le situer et l'afficher. Parmi les objets de l'univers, nous avons ressenti le besoin d'introduire des objets spatiaux (*'f_spatial'*) afin de permettre à l'utilisateur de manipuler éventuellement autre chose que des zones temporelles ; ce point est de peu d'importance pour l'instant. Par contre, la décomposition des zones en trois catégories : *'f_attention'*, *'f_texte'* et *'f_prédicat'* donne à chaque instance de ces différentes classes un aspect graphique différent facilitant leur manipulation à l'écran ; leur comportement comme zones temporelles n'en est cependant pas affecté.

Les deux variables 'objet_1' et 'objet_2' du flavor 'f_relation' contiennent les deux objets reliés par une instance de ce flavor (ou de sa descendance). Un double lien est donc établi entre les objets et les relations, permettant à tout moment de passer de l'un à l'autre et réciproquement par un minimum de calcul.

Maintenant que nous possédons les principaux éléments de notre système, nous pouvons en donner le schéma de fonctionnement général tel qu'il est résumé dans la figure 4.2. La conscience, qui dirige d'un point de vue fonctionnel l'ensemble du système, est traversée par deux flux de données transversaux : des données textuelles et graphiques. Ces deux parties sont à la base complètement dissociées, du fait de leur nature conceptuelle différente. D'un coté, les informations textuelles concernent le système en tant qu'implémentation de notre modèle, de l'autre, l'interface graphique n'influe en rien sur le comportement interne des mécanismes du modèle.

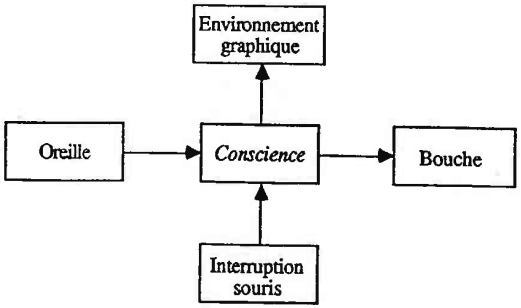


Figure 4.2 : schéma de fonctionnement du système.

4.3 FONCTIONNALITÉS DISPONIBLES.

4.3.1 Gestion des interfaces textuelles.

Peu de zones sont instanciées par avance lors du lancement du système, parmi celles-ci une instance de 'f_bouche' et une de 'f_oreille' gèrent les échanges entre l'utilisateur et la conscience. Pour l'instant, l'oreille transforme, dans un sens, une suite de caractères en une zone de type 'f_texte' qu'elle transmet à la conscience et dans l'autre, la bouche transforme une zone en un message à l'utilisateur, en fonction de la valeur de la variable 'texte'.

Suivant le type des informations que l'on désire manipuler, ces deux interfaces peuvent être compliquées à souhait. En particulier, le lien avec les systèmes de décodage acoustico-phonétique dont nous disposons actuellement au Crin est immédiat à réaliser, puisqu'un treillis phonétique n'est rien d'autre dans notre modèle qu'un ensemble de zones temporelles reliées entre elles par des relations d'adjacence. On peut alors imaginer qu'un tel treillis soit transmis à la conscience au fur et à mesure de sa création.

4.3.2 Interface graphique.

1 AFFICHAGE DES ZONES.

L'interface graphique entre le système et l'utilisateur passe par une fenêtre où apparaissent certains objets actifs à un certain stade de l'analyse. A chaque mise à jour de l'affichage dans cette fenêtre, toutes les zones présentes dans la conscience sont rendues

visibles, ainsi que les zones qui leur sont reliées et qui sont marquées comme actives. Pour chaque zone, on explore ainsi les relations d'inclusion, d'adjacence et d'héritage, et ceci récursivement jusqu'à aboutir sur une zone non-active et qui n'est pas dans la variable 'attention'. Les différentes relations ne sont tracées que lorsque leurs deux extrémités sont déjà visibles à l'affichage, afin d'éviter la présence de relations "en l'air" dont on ne pourrait déterminer la destination.

La mise en place de l'affichage des différentes zones nécessite la gestion de la topologie de la fenêtre, afin qu'aucun chevauchement n'apparaisse et qu'une certaine cohérence de tracé soit préservée par rapport à la sémantique des relations. Ainsi, pour chaque nouvel objet à tracer, une recherche rapide d'un espace de tracé libre est effectuée, en fonction de la relation temporelle ou d'héritage qui a permis d'accéder à cet objet. Ainsi, si une relation d'adjacence est à l'origine de l'affichage d'une zone, la recherche d'un emplacement se fera horizontalement à droite ou à gauche suivant les cas. De même, pour les relations d'inclusion ou d'héritage, le parcours se fait verticalement.

2 DÉSIGNATION DE ZONES.

Les objets de l'univers qui ont été visualisés peuvent être désignés directement par l'intermédiaire de la souris. Dès lors, cette opération est signalée à la conscience qui mémorise au fur et à mesure ces objets, jusqu'à ce que l'on obtienne une paire d'objets différents qui peuvent être appariés. Dans ce cas un menu est proposé à l'utilisateur qui choisit entre les opérations d'instanciation ou d'abstraction en cas de réussite de l'appariement.

Ainsi, il n'y a pas pour l'instant de mode automatique de fonctionnement dans lequel la conscience gérerait d'elle-même les objets qu'elle manipule. L'interface graphique présente alors l'avantage pour l'utilisateur d'un minimum d'opérations manuelles à effectuer pour faire fonctionner le système.

4.3.3 Gestion de la conscience.

La conscience fonctionne à partir d'une boucle infinie qui traite en continu les différentes informations qui lui parviennent et effectue les opérations adéquates en fonction de la nature de celles-ci. Ces opérations peuvent être de différents ordres :

- insertion d'une nouvelle zone envoyée par l'oreille à la suite de la frappe d'un message par l'utilisateur.

- réponse à une commande particulière (par l'intermédiaire de l'oreille toujours), à savoir : inhibition de l'affichage, remise à jour de celui-ci, opérations sur les zones, sauvegarde ou récupération d'un ensemble de zones.

- réponse à une interruption souris. Dans ce cas elle lance l'appariement de deux zones dès que celles-ci sont détectées.

Dès la mise en place d'un mode de fonctionnement automatique, la conscience sera chargée de faire évoluer son attention en fonction des possibilités de déduction qui sont offertes sur les zones actives.

4.3.4 Appariement de zones.

Le mécanisme d'appariement suit exactement la démarche qui a été envisagée au troisième chapitre. La recherche s'effectue récursivement en fonction d'un contexte d'appariement (liste de couples déjà appariés) qui est construit à chaque étape. L'ensemble repose en alternance sur deux phases :

- appariement de deux relations : cette opération n'est possible que si les deux relations sont identiques et parcourues dans le même sens, ou bien si une des deux est vide (objet : '$vide'), c'est à dire lorsqu'on est en train de parcourir un sous-ensemble de zones ne possédant pas de pendant dans l'autre groupement à apparier. L'appariement peut aussi être rejeté si une des deux relations est déjà présente dans le contexte local d'analyse. Le résultat de cette opération est alors un nouveau couple de zones à étudier correspondant aux extrémités respectives des deux relations.

- appariement de deux zones temporelles : les deux zones doivent posséder des ancêtres communs au sens de la relation d'héritage et ne pas avoir été associées à d'autres zones au cours des appariements précédents (vérification du contexte d'appariement). Eventuellement, un des objets peut avoir déjà été associé à un objet vide ; il faut alors mettre à jour le contexte d'analyse. Une fois vérifié l'appariement possible entre les deux zones, on génère un ensemble de nouveaux choix d'analyse de la façon suivante : si R1 est l'ensemble des relations temporelles partant du premier objet en cours d'analyse, et R2 l'ensemble des relations partant du deuxième, on calcule toutes les mises en correspondance possibles de R1 ∪ { $vide } et de R2 ∪ { $vide } à l'exclusion du couple ($vide , $vide). On obtient ainsi autant de possibilités d'analyse qui seront toutes étudiées, à moins qu'elles ne conduisent à des impasses.

Donc, le résultat de cette opération est formé d'un ensemble de contextes pour toutes les analyses qui ont réussi pour un parcours complet des voisins des deux zones initiales à étudier. Ce résultat sera utilisé suivant les cas, soit pour une instanciation de zones, soit pour faire une abstraction le long de la relation d'héritage, après une éventuelle vérification de la cohérence.

4.3.5 Vérification de la cohérence temporelle.

A partir du résultat d'un appariement de deux ensembles de zones, on extrait l'ensemble des arcs correspondants (les couples de relations) sur lesquels on applique l'algorithme présenté au 3.5.2. Conformément à celui-ci, le résultat de la procédure est soit une valeur d'erreur, soit la liste d'origine complétée d'éventuelles déductions temporelles.

4.3.6 Abstraction/Instanciation de zones.

Les différentes possibilités présentées au 3.5.3 ont été entièrement mises en place dans le système actuel. La seule restriction concerne le choix des associations que l'on instancie effectivement quand un résultat multiple est obtenu lors de l'étape d'appariement. Ce problème est encore à l'heure actuelle délicat à résoudre, puisqu'il influencera en particulier les capacités d'apprentissage du système. Notre opinion est qu'il est important de conserver un maximum de ces appariements pour les laisser disparaître d'eux-mêmes s'ils n'interviennent pas ultérieurement dans d'autres analyses.

4.3.7 Sauvegarde de zones temporelles.

La possibilité d'enlever temporairement certaines parties de l'espace de représentation du système est fondamentale pour une amélioration de la qualité des traitements. En effet, cela permet de limiter à la fois la place occupée en mémoire et le nombre d'objets manipulés à un instant donné par les mécanismes de raisonnement. Nous avons donc mis en place deux opérations complémentaires de préservation de l'environnement :

- sauvegarde d'une zone, ainsi que de tous ses voisins au sens de l'adjacence et de l'inclusion. Cette opération peut être limitée au contenu d'une zone (de cohérence) donnée, afin de localiser son effet.

- restitution à partir d'un fichier d'un ensemble de zones faisant référence à des zones déjà en mémoire qui servent alors de point de rattachement.

La sauvegarde s'effectue par parcours récursif des relations d'adjacence et d'inclusion. Pour chaque objet ou relation rencontré, on génère un nouveau symbole qui sert de référence dans le fichier final. La description des différentes variables internes (hormis celles qui concernent le graphique) est faite de la même manière, en remplaçant chaque objet par son équivalent symbolique. Le résultat est donc obtenu sous la forme d'un fichier texte, qu'il est alors possible d'éditer pour le modifier ou le compléter. La définition d'un lexique de base, accompagné de certaines structures syntaxiques utiles pour les tests en cours en est une application immédiate.

4.4 Compléments envisagés.

Deux compléments principaux sont en cours de développement afin d'accroître la puissance de cet atelier de manipulation de zones temporelles :

- dans un premier temps, nous nous attachons à renforcer l'importance des zones de cohérence dans les différentes fonctionnalités déjà étudiées. En particulier, il est nécessaire de restreindre l'appariement de deux ensembles de zones aux contenus de deux zones de cohérence particulières, pour éviter une explosion combinatoire importante si on explore un trop grand nombre de zones. Cette restriction est aussi indispensable lors de la sauvegarde d'un ensemble de zones, pour les mêmes raisons. Ce travail doit s'accompagner d'une réflexion en profondeur relative au statut des zones de cohérences qui sont bien plus que de simple zones temporelles dès lors qu'elle déterminent des espaces clos de raisonnement. C'est à l'heure actuelle un de nos objectifs prioritaires que de maîtriser cet aspect de notre modèle.

- un étape importante pour aboutir à un système opérationnel consiste à passer du fonctionnement semi-automatique actuel à un fonctionnement où la conscience choisit d'elle-même les objets sur lesquels elle travaille. Là encore, en dehors d'un développement théorique plus important, on ne pourra s'appuyer que sur un ensemble de stratégie relativement primaire. Cet automatisation doit alors s'accompagner d'une interface de définition de zones plus pratique que l'utilisation d'un fichier texte, comme cela se fait actuellement, ainsi que d'un mécanisme destiné à proposer diverses stratégies de recherche le long de l'arborescence.

On constate donc, au vu de ce court exposé relatif à l'implémentation de notre modèle de représentation des informations temporelles, que nous attachons une grande importance à ne mettre en place que les aspects de ce dernier dont les bases conceptuelles sont claires. En effet, nous pensons qu'il est inutile de développer des programmes en tâtonnant pour aboutir finalement à un résultat hasardeux. Cependant, l'implémentation permet à chaque fois de mettre le doigt sur des problèmes qui n'apparaissent pas sur le papier et c'est donc un complément indispensable.

5 Conclusions.

Les recherches présentées tout au long de cette thèse ont visé dans leur ensemble à définir de nouvelles approches pour l'étude du dialogue homme-machine. Nous sommes encore loin, à l'heure actuelle, de la définition complète d'un système autonome qui dialoguerait de façon naturelle avec un utilisateur. Pourtant, nous sommes partis de beaucoup plus loin encore lorsque nous avons commencé à définir un gestionnaire d'hypothèses lexicales dans le cadre d'une architecture de système modulaire et figée (chapitre 1). Le constat de semi-échec auquel nous avons abouti à la suite de cette étude a guidé nos recherches vers d'autres domaines d'analyse, dont les sciences cognitives, afin de répondre à un problème important : *est-il possible, dans le cadre d'un modèle de représentation unique, d'intégrer toutes les connaissances nécessaires à un dialogue homme-machine ?*

Dans le deuxième chapitre, nous avons ainsi parcouru de nombreuses propositions touchant différents aspects du dialogue, de la langue à l'univers du discours. Nous les avons jugées au regard de nos objectifs, ce qui a permis de dégager certaines grandes orientations. Parmi celles-ci, nous avons plus particulièrement remarqué la nécessité de *modéliser l'activité du locuteur/auditeur en situation de communication*, plutôt que d'envisager la langue (et les informations qu'elle transmet) indépendamment de l'usage qui peut en être fait par un individu.

A partir de ces résultats, nous avons proposé un axe d'étude réduit au seul domaine des informations temporelles, qui malgré tout, recoupe les différents niveaux de traitement couramment envisagés pour le langage. C'est ainsi que nous avons été amenés dans le chapitre trois à définir un *modèle cognitif de représentation des informations temporelles* qui tente de répondre aux besoins de *représenter, apprendre et prédire* ; trois points cruciaux pour un dialogue *naturel*. Ce modèle, simplifié par rapport aux autres représentations du temps, repose sur un seul type d'objet : les zones temporelles, ainsi que sur deux relations fondamentales : l'adjacence et l'inclusion de deux zones. A ces objets nous avons ajouté un lien d'héritage entre les zones qui permet à notre modèle d'apprendre et de prédire des structures temporelles particulières.

Bien que ce modèle temporel réponde dans le cadre restreint que nous nous étions imposé à tous nos objectifs, nous n'avons pas pour autant résolu le problème du dialogue homme-machine dans son ensemble. Notre modèle est d'ailleurs apparu, en conclusion

du troisième chapitre, bien plus comme un espace de recherche ouvert sur de futurs développements que comme un bloc monolithique dont nous ne voudrions pas voir modifier un seul élément. C'est pourquoi nous envisageons d'ores et déjà différents objectifs de recherche à partir des travaux que nous avons présentés.

Nous devons tout d'abord mettre en œuvre totalement l'atelier de manipulation de zones temporelles défini au chapitre quatre, afin de valider notre modèle dans différents domaines d'applications :

- *le dialogue oral homme-machine*, cadre initial de ce travail.
- *la compréhension de récit*, pour automatiser le traitement des exemples étudiés au cours du troisième chapitre, en nous limitant toujours aux informations temporelles.
- *la génération de plan*, car c'est un domaine privilégié où le temps intervient et pour lequel notre modèle semble bien adapté.

Sur un plan plus théorique, nous avons déjà observé certaines lacunes de notre modèle, qu'il serait nécessaire de combler dans un avenir relativement proche. Parmi celles-ci, le problème des possibles, très important par ailleurs en psychologie ([Piaget 81]), a été rencontré en différents endroits lors de l'analyse du présent rapporté, du futur ou des prédictions multiples. En tenir compte donnerait à ce modèle une dimension supplémentaire.

Un autre aspect concerne l'extension de notre modèle à d'autres aspects parallèles dont la représentation des objets spatiaux fait partie. Les deux domaines du temps et de l'espace semblent en effet très liés, comme le montrent les expressions linguistiques par lesquelles ils sont exprimés. Il semble que ceci traduise une analogie plus profonde dont il faudrait tenir compte.

Enfin, nous conservons notre but principal, même s'il est de plus longue haleine, à savoir la *définition de nouvelles architectures de systèmes de dialogue homme-machine* en prolongement de nos recherches actuelles.

En conclusion, nous pouvons souligner combien notre approche s'écarte résolument des représentations mathématiques (pour le temps) ou logique (pour le raisonnement). Nous avons essayé de montrer qu'il existe une alternative *cognitive* à des théories parfois un peu trop formelles, à condition de se situer dans certains cadres d'analyse où l'aspect humain importe plus que la précision des opérations effectuées.

6 BIBLIOGRAPHIE.

[Aho 72] V.A. Aho et J.D. Ullman, *The theory of parsing*, Prentice Hall, EnglewoodCliffs, N.J., 1972.

[Alegria 83] J. Alegria, *L'espace et le temps aujourd'hui*, Seuil, Paris, 1983.

[Alexandre 88] F. Alexandre, Y. Burnot, F. Guyot, et J.P. Haton, "La colonne corticale : nouvelle unité de base pour les réseaux multi-couches", *Actes de Neuro-Nimes'88*, Nimes, 15-17 Nov., 1988, pp.21-33.

[Alinat 87] P. Alinat, E. Gallais, J.P. Haton, J.M. Pierrel et P. Richard, "A continuous speech dialog system for oral control of sonar console", Proc. *IEEE ICASSP-87*, 1987.

[Allen 83] James F. Allen, "Maintaining knowledge about temporal intervals", *Communications of the ACM*, Vol.26 n°11, 1983, pp.832-843.

[Allen 84] James F. Allen, "Towards a general theory of action and time", *A.I. Journal*, pp.123-154.

[Allen 85] James F. Allen et Patrick J. Hayes, "A common-Sense Theory of time", *Proc. IJCAI 85*, 1985, pp.528-531.

[Amalberti 88] R. Amalberti, N. Carbonell et P. Falzon, "Dialogue Homme-Homme, Dialogue Homme-Machine : Un même modèle ?", *Actes du 3ème Colloque International de l'ARC*, Toulouse, 9-11 mars, 1988, pp.231-246.

[Anderson 80] John R. Anderson, *Cognitive Psychology*, W. H. Freeman and Cie, New York, 1980, 2nd Edition 1985.

[Aouati 84] A. Aouati, D. Béroule, J.-J. Mariani et F. Néel, "Différentes approches du dialogue homme-machine au LIMSI", in [Pierrel 84], pp.61-90.

[ARC 88] Association pour la Recherche Cognitive, *Actes du 3ème colloque international, Cognition et Connaissance : où va la science cognitive*, Toulouse, 9-11 mars, 1988.

[Barnett 83] J.A. Barnett, "Computational methods for a mathematical theory of evidence", *Proc. 8th IJCAI*, Karlsruhe, W.G., Août 8-12, 1983, pp. 868-875.

[Béroule 84] D. Béroule et F. Néel, "Une approche des problèmes liés à la communication parlée homme-machine", *Actes du 4ème Cong. AFCET-RFIA*, Paris, 1984, pp. 53-63.

[Béroule 85] D. Béroule, *Un modèle de mémoire adaptative, dynamique et associative pour le Traitement Automatique de la Parole*, Thèse de 3ème cycle, Orsay, 1985.

[Béroule 88a] D. Béroule, "The adaptive, dynamic and associative memory model : a possible future tool for vocal human-computer communication", *The structure of multimodal dialogue*, M.M.Taylor, F.Néel et D.G.Bouwhuis (eds), North Holland, Amsterdam, 1988.

[Béroule 88b] D. Béroule, "The never-ending learning", *Neural computers*, R. Eckwiller, C.v.d. Malsburg (eds), Springer Verlag, 1988.

[Bertille 88] J.M. Bertille et J.C. Perez, "Le modèle holographique «chaos fractal» : bases théoriques et applications", *Actes de Neuro-Nimes'88*, Nimes, 15-17 Nov., 1988, pp.263-282.

[Bonté 88] E. Bonté, J. Castaing, P. Grandemange, S. Grumbach, D. Kayser et F. Lévy, *Description générale d'un raisonneur à profondeur variable*, Rapport L.I.P.N., 1988.

[Borillo 86] Andrée Borillo, "Les emplois adverbiaux des noms de temps", *Actes du séminaire «Lexique et traitement automatique des langages»*, Univ. P. Sabatier, Toulouse, 16-17 Janv., 1986.

[Borillo 88] Andrée Borillo et Mario Borillo, "Une approche cognitive du raisonnement temporel", *Actes des journées internationales du PRC-GRECO Intelligence Artificielle*, Toulouse, 14-15 mars, 1988, pp.11-36.

[Bornerand 88] S. Bornerand, F. Néel, G. Sabah, "Un modèle de langage unifié dans un système de dialogue oral pilote/avion", *Actes des XVII^èmes J.E.P.*, Nancy, 20-22 Sept., 1988, pp.61-64.

[Bourdieu 82] Pierre Bourdieu, *Ce que parler veut dire*, Fayard, Paris, 1982.

[Bundy 85] Alan Bundy, "Incidence Calculus: A mechanism for Probabilistic Reasoning", *Journal of automated Reasoning*, Vol.1, 1985.

[Bundy 86] Alan Bundy, "Correctness criteria of some algorithms for uncertain reasoning using incidence calculus", *Journal of automated Reasoning*, Vol.2, 1986.

[Burnod 87] Y. Burnod, *Cerebral cortex and behavioral adaptation : a possible mechanism*, Presses universitaires de Liège, Bel., 1987.

[Carbonell 84] N. Carbonell, F. Charpillet, J.P. Haton, B. Mangeol, P. Mousel, J.M. Pierrel, A. Roussanaly, "Dialogue oral homme-machine : Bilan du projet myrtille et perspectives", in : *Actes du Séminaire GRECO*, 11-12 octobre, 1984.

[Carbonell 87] N. Carbonell et J.M. Pierrel, "Architecture of knowledge sources in a human-computer oral dialogue system". in : M.M. Taylor, F. Neel et D.G.Bouwhuis, eds., *structure of multimodal dialogues*, North-Holland, Amsterdam, 1987 .

[Carbonell 88] N. Carbonell et J.J. Bonin, "Détection de frontières syntagmatiques en parole continue : utilisation de la fréquence fondamentale", *Actes des XVII^èmes J.E.P.*, Nancy, 20-22 Sept., 1988, pp.163-167.

[Chailloux 86] J. Chailloux, *Le_lisp 15.2, Le manuel de reference*, INRIA, 1986.

[Changeux 83] Jean Pierre Changeux, *L'homme neuronal*, Fayard, Paris, 1983.

[Cheeseman 85] P. Cheeseman, "In Defense of Probability", *Proc. 9th IJCAI*, Los Angeles, Cal., Août 18-23, 1985, pp. 1002-1009.

[Chomsky 65] N. Chomsky, *Aspects de la théorie syntaxique*, Le seuil, Paris, 1971, traduction de *Aspects of the theory of syntax*, MIT Press, 1965.

[Cointe 84] Pierre Cointe, "Formes par l'exemple ou de la mise en Formes", *Actes de BIGRE n°41*, AFCET, Brest, nov., 1984.

[Cottrell 84] Garrison W. Cottrell, "A model of lexical access of ambiguous words", *AAAI-84*, p61-p67.

[Danlos 88] L. Danlos, "Les problèmes posés par les verbes supports en traduction automatique", in : *actes du colloque Informatique et Linguistique*, Nantes, 12-13 oct., 1988, à insérer p. 255.

[De Mori 83] R. De Mori, *Computer models of speech using fuzzy algorithms*, Plenum Press, New York, 1983.

[den Heyer 85] Ken den Heyer et Annette Goring, "Semantic Priming and Word Repetition : The Two Effects Are Additive", *Journal of Memory and Language*, Vol.24, 1985, pp.699-716.

[Dennett 81] Daniel C. Dennett, *Brainstorms*, Harvester Press, Brighton, G.B., 1981.

[Dennett 87] Daniel C. Dennett, *The intentional stance*, M.I.T. press, Cambridge, Mass., 1987.

[Deville 86] G. Deville et H. Paulussen, *A case grammar as an original linguistic model for the semantic representation of utterances in a man-machine dialogue system*, post-graduate Thesis, Antwerpen, Belgium, Sept. 1986.

[Deville 87a] G. Deville, H. Paulussen et J.M. Pierrel, "Une grammaire de cas comme modèle de représentation sémantique d'énoncés de dialogues oraux homme-machine finalisés", *Actes du 6ème Cong. AFCET-RFIA*, Antibes, Nov. 1987, pp. 157-174.

[Deville 87b] G. Deville, *Corpus de dialogues oraux finalisés en situation réelle : Demande de renseignements auprès du ministère de l'emploi et du travail (Cellule Action-travail)*, FUNDP Namur, 1987.

[Deville 89] G. Deville, "Modelization of task-oriented utterances in a man-machine dialogue system", *Thesis of Doctor of Linguitics*, University Instelligen, Antwerpen, 1989.

[Dubois 80] D. Dubois et H. Prade, *Fuzzy Sets and Systems, Theory and Applications*, Academic Press, 1980.

[Dubois 86] D. Dubois et H. Prade, *"Théorie des possibilités. Applications à la représentation des connaissances en informatique"*, Masson 86.

[Dubois 88] Danièle Dubois et Liliane Sprenger-Charolles, "Perception/Interprétation du langage écrit : contexte et identification des mots en cours de lecture", *Intellectica*, n°5, 1988, pp.113-146.

[Eco 65] U. Eco, *L'œuvre ouverte*, Seuil, Paris, 1965, réédition : Points, Seuil, 1979.

[Eco 72] U. Eco, *La structure absente*, Mercure de France, Paris, 1972.

[Eco 85] Umberto Eco, *Lector in Fabula*, Grasset, Paris, 1985.

[Erman 80a] L.D. Erman, F. Hayes-Roth, V.R. Lesser et D.R. Reddy, "The Hearsay II Speech Understanding System, Integrating Knowledge to Resolve Uncertainty", *Computing Surveys*, Vol.12, n°2, June 1980.

[Erman 80b] Lee D. Erman et Victor R. Lesser, "The Hearsay-II Speech Understanding System : A Tutorial". in : W.A. Lea, *Trends in Speech Recognition*, Prentice-Hall, 1980.

[Fauconnier 85] Gilles Fauconnier, *Mental Spaces*, M.I.T. Press, Cambridge, Mass., 1985.

[Fikes 71] R.E. Fikes et N.J. Nilsson, "STRIPS : A new approach to the application of theorem proving to problem solving", *AIJournal*, Vol.2, 1971.

[Fillmore 68] C. Fillmore, "The case for the case", in *Universals in linguistic theory*, Bach et Harms, Chicago, Holt, Rinehart and Winston, pp.1-90.

[Flieller 86] André Flieller, *La coéducation de l'intelligence*, Presses Universitaires de Nancy, 1986.

[Fogelman 88] Françoise Fogelman-Soulie, "Méthodes connexionnistes pour l'apprentissage", in : *Actes des journées nationales du PRC-GRECO I.A.*, Tulouse, 14-15 mars, 1988, pp.275-293.

[Fohr 87] D. Fohr, N. Carbonel, J.P. Haton, "APHODEX, an acoustic-phonetic decoding expert system", *International Journal of Pattern Recognition and Artificial Intelligence*, C.H.Chen ed., Vol.1 n°2 1987, pp. 207-222.

[Fuchs 87] Catherine Fuchs, ed, *L'ambiguïté et la paraphrase*, Université de Caen, 1987.

[Garvey 83] Thomas D. Garvey, John D. Lowrance et Martin A. Fischler, "An inference technique for integrating knowledge from disparate sources", *IJCAI 83*.

[Ginsberg 84] M.L. Ginsberg, "Non-monotonic reasoning using Dempster's rule", *Proc. AAAI 84*, 1984, pp. 126-129.

[Ginsberg 85] Matthew L. Ginsberg, "Does Probability have a place in non-monotonic reasoning", *IJCAI 85*, p107-p110.

[Girard 87] J.-Y. Girard, "Linear Logic", *Theorical Computer Science*, n°50, North Holland, 1987, pp.1-102.

[Gordon 85] Barry Gordon, "Subjective Frequency and the Lexical Decision Latency Function : Implications for Mechanisms of Lexical Access", *Journal of Memory and Language*, Vol.24, 1985, pp.631-645.

[Granier 88] T. Granier, *Contribution à l'étude du temps objectif dans le raisonnement*, rapport de recherche RR 716-I-73 LIFIA, Grenoble, 1988.

[Gross 68] Maurice Gross, *Grammaire transformationnelle du français : T.1 Syntaxe du verbe*, Cantilène, Paris, 1968, 1986.

[Gross 77] Maurice Gross, *Grammaire transformationnelle du français : T.2 Syntaxe du nom*, Cantilène, Paris, 1968, 1986.

[Gross 88] M. Gross, "Dictionnaire électronique et reconnaissance des mots dans les textes", in : *actes du colloque Informatique et Linguistique*, Nantes, 12-13 oct., 1988, pp.259-288.

[Hagège 76] Claude Hagège, *La grammaire générative - réflexions critiques*, PUF, Paris, 1976.

[Hagège 85] Claude Hagège, *L'homme de paroles*, Fayard, Coll. Le temps des sciences, Paris, 1985.

[Hart 78] P.E. Hart, R.O. Duda et M.T. Einaudi, "A Computer-based Consultation System for Mineral Exploration", Technical report, SRI International, May 1978.

[Haton 77] J.P. Haton, J.F. Mari et J.M. Pierrel, "Research towards speech understanding models for artificial and natural language", *Proc. IEEE ICASSP-77*, Dallas, Texas, 1977, pp.807-810.

[Hayes 87] Patrick J. Hayes et James F. Allen, "Short Time Periods", *Proc. IJCAI 87*, 1987, pp.981-983.

[Heyer 85] K. den Heyer, A. Goring et G.L. Dannenbring, "Semantic priming and word repetition, the two effects are additive", *Journal of memory and langage*, Vol.24, 1985, pp.699-716.

[Hornstein 77] Norbert Hornstein, "Towards a Theory of tense", *Linguistic Inquiry*, Vol.8 n°3, 1977.

[Hudson 84] Richard Hudson, *Word Grammar*, Blackwell, Oxford, G.B., 1984.

[Kaufman 73/77] Arnold Kaufman, *"Introduction à la théorie des sous-ensembles flous"*, Tomes 1-4, Masson, 1973, 1975, 1977.

[Kaufman 87] Arnold Kaufman, *"Nouvelles logiques pour l'I.A."*, Hermes, 1987.

[Kayser 87] D. Kayser, "Raisonnement : logique et informatique", *Intellectica*, Vol.1 n°4, pp.81-103, 1987.

[Kayser 88] Daniel Kayser, "Le raisonnement à profondeur variable", *Actes des journées nationales du PRC-GRECO IA*, Toulouse, 14-15 mars, 1988.

[Kosslyn 78] S.M. Kosslyn, T.M. Ball et B.J. Reiser, "Visual images preserve metric spatial information : evidence from studies of image scanning", *Journal of experimental psychology : human perception and performance*, Vol.4, 1978, pp.47-60.

[Kumar 87] Krishna Kumar et Amithaba Mukerjee, "Temporal Event Conceptualization", *Proc. IJCAI 87*; 1987, pp.472-475.

[Ladkin 87] P. Ladkin, "The completeness of a natural system for reasoning with time intervals", *Proc. IJCAI 87*; 1987, pp.462-467.

[Laprie 88] Y. laprie, "Snorri, Un système d'étude interactif de la parole", *Actes des XVII*èmes *J.E.P.*, Nancy, 20-22 Sept., 1988, pp.71-76.

[Latraverse 87] François Latraverse, *La pragmatique : histoire et critique*, P. Mardaga, Bruxelles, 1987.

[Lâasri 88] H. Lâasri, B. Maître, F. Charpillet, T. Mondot et J.P.Haton, "ATOME : A Blackboard Architecture with Temporal and Hypothetical Reasoning", *Proc. 8th ECAI*, Munich, Août 1-5, 1988, pp.5-10.

[Le Cun 87] Yann Le Cun and Françoise Fogelman-Soulié, "Modèles connextionnistes de l'apprentissage", in Joël Quinqueton and Jean Sallantin eds. *intellectica v1 n2/3, apprentissage et machine.*

[Le Ny 79] Jean-François Le Ny, *La sémantique psychologique*, P.U.F., Paris, 1979.

[Lea 80] W.A. Lea, *Trends in Speech Recognition*, Prentice-Hall, 1980.

[Levinson 83] Stephen C. Levinson, *Pragmatics*, Cambridge University Press, Cambridge, 1983.

[Lowerre 80] B. Lowerre et R. Reddy, "The Harpy Speech understanding system", in : W. A. Lea, eds.,*Trends in speech recognition*, Prentice Hall 1980, pp. 340-360.

[Lu 84] S. Y. Lu et H. E. Stephanou, "A set-theoretic framework for the processing of uncertain knowledge", *Proc. AAAI 84*, pp. 216-221.

[Lyons 77] John Lyons, *Semantics*, Cambridge University Press, Cambridge, 1977, tome 1-2.

[Mariani 78] J.-J. Mariani et J.-S. Liénard, "ESOPE 0 : un programme de compréhension automatique de la parole", in : *Proc. Congrès AFCET-RFIA*, Paris 1978, pp. 169-175.

[Marslen-Wilson 80] W.D. Marslen-Wilson et L.K. Tyler, "The temporal structure of spoken language understanding", *Cognition*, Vol.8, 1980, pp.1-71.

[Martin 83] R. Martin, *Pour une logique du sens*, PUF, Paris, 1983.

[McCarthy 86] John McCarthy, "Applications of circumscription to formalizing Common-sense knowledge", *AIJournal*, Vol.28 n°1, pp.89-116, 1986.

[McDermott 82] D. McDermott, "A temporal logic for reasoning about processes and plans". *Cognitive Science*, Vol.6, 1982, pp.101-155.

[Merialdo 86] B. Merialdo, A.M. Derouault et S. Soudoplatoff, "Phoneme classification using Markov models", *Proc. ICASSP 86*, Tokyo, 1986.

[Minsky 75] M. Minsky, "A Framework for Representing Knowledge", in *The psychology of computer vision*, P. Winston (Ed.), McGraw-Hill, New York, 1975.

[Moeschler 85] J. Moeschler, *Argumentation et conversation. Eléments pour une analyse pragmatique du discours*, Hatier, 1985.

[Moon 86] D.A. Moon, "Object Oriented Programming with Flavors", *Proc. OOPSLA's 86*, Portland, Or., Sept.29-Oct.2, 1986, pp. 1-8.

[Morin 87] P. Morin et J.M. Pierrel, "PARTNER : un système de dialogue oral homme-machine", *Actes du congrès COGNITIVA*, Paris, 1987, pp.354_361.

[Moulin 88] B. Moulin et A. Kabbaj, "Architecture de SMGC, un système de manipulation de graphes conceptuels", in : *Actes du colloque Informatique et Linguistique*, Nantes, 12-13 oct., 1988, pp.71-98.

[Nef 84] Frédéric Nef, *Sémantique de la référence temporelle en français moderne*, Peter Lang, Bern, 1984.

[Nef 88] F. Nef, *Logique et Langage*, Hermes, Paris, 1988.

[Nilsson 80] Nils J. Nilsson, *Principles of Artificial Intelligence*, Springer-Verlag, Berlin, 1980.

[Oswald 86] Jacques Oswald, *Théorie de l'information ou Analyse Diacritique des systèmes*, CNET/ENST, Masson, Paris, 1986.

[Papoulis 65] Athanasios Papoulis, *Probability, Random Variables and Stochastic Processes*, McGraw-Hill, 1965, 1984.

[Piaget 46] Jean Piaget, *Le developpement de la notion de temps chez l'enfant*, P.U.F., Paris, 1946, 3ème ed., 1981.

[Piaget 79] Jean Piaget et Noam Chomsky, *Théories du langage et théories de l'apprentissage*, Seuil, Paris, 1979.

[Piaget 81] Jean Piaget, *Le possible et le nécessaire*, PUF, Paris, 1981.

[Pierrel 78] J.M. Pierrel, "Un système de compréhension automatique du discours continu utilisant des contraintes morphologiques, syntaxiques et sémantiques", *RAIRO Informatique*, Vol.12-2, pp.109-120, 1978.

[Pierrel 79] J.M. Pierrel et J.P. Haton, "Data Structure and Organization of Myrtille II System", *Proc. 4th IJCPR*, Kyoto, 1979.

[Pierrel 81] J.M. Pierrel, *Etude et mise en œuvre de contraintes linguistiques en compréhension automatique du discours continu*, Thèse d'état, Université de Nancy I, 1981.

[Pierrel 82] J.M. Pierrel, *"Utilisation de contraintes linguistiques en compréhention automatique de la parole continue"*, *T.S.I.*, Vol.1-5, 1982, pp. 403-421.

[Pierrel 84] J.M. Pierrel, N. Carbonell, J.-P. Haton et F. Néel (eds), *Dialogue homme-machine à composante orale*, GRECO-CNRS Communication parlée, Nancy, 442p.

[Pierrel 87] Jean-Marie Pierrel, *Dialogue Oral Homme-Machine, connaissances linguistiques, stratégies et architectures des systèmes*, Hermes, Paris, 1987.

[Prade 82] H. Prade, *Modèles mathématiques de l'imprécis et de l'incertain en vue d'applications au raisonement naturel*, Thèse d'état, Université P. Sabatier, Toulouse, 1982.

[Quillian 68] R. Quillian, "Semantic Memory", in : *Semantic Information Processing*, M. Minsky ed., M.I.T. Press, Cambridge, Mass., 1968.

[Reichenbach 47] Reichenbach, *Elements of symbolic logic*, 1947.

[Reichman 86] Rachel Reichman, *Getting Computer to talk like you and me*, M.I.T. Press, Cambridge, Mass., 1986.

[Reiter 80] Raymond Reiter, "A logic for defauly reasoning", in *AIJournal*, Vol.13 n°1-2, 1980, pp.81-132.

[Rich 83] Elaine Rich, *Artificial Intelligence*, McGraw-Hill, 1983.

[Romary 87] L. Romary, *Quelques problèmes liès à l'incertain en reconnaissance de la parole*, rapport de D.E.A., Université de Nancy I, Juin, 1987.

[Romary 88] L. Romary et B. Mangeol, "Prédiction et vérification lexicale dans le cadre d'un dialogue oral homme-machine", *Proc. XVII*ème *J.E.P.*, Nancy, Sept. 20-22, 1988, pp.168-172.

[Romary 89] L. Romary et J.M. Pierrel, "The use of Dempster-Shafer rule in the lexical component of a man-machine oral dialogue system", *Speech Communication*, à paraître, 1989.

[Rose 75] Steven Rose, *Le cerveau conscient*, Editions du Seuil, Paris, 1975.

[Roussanaly 86] A. Roussanaly, P. Mousel, N. Carbonnell, B. Mangeol et J.M. Pierrel, *Réalisation d'un corpus de dialogues oraux. Application aux renseignements administratifs*, Rapport CRIN n°86-R-083, 1986.

[Roussanaly 88] A. Roussanaly, *DIAL : la composante dialogue d'un système de communication orale homme-machine finalisée en langage naturel*, Thèse de doctorat de l'université de Nancy I, 1988.

[Rumelhart 86] David R. Rumelhart, James L. McClelland and the PDP research group, *Parallel distributed processing*, T.1-2, M.I.T. Press, Cambridge, 1986.

[Russell 59] Bertrand Russell, *Signification et vérité*, Flammarion, Paris, 1959.

[Sabah 88] G. Sabah, *L'intelligence artificielle et le langage*, Hermes, Paris, 1988.

[Sacerdoti 74] E.D. Sacerdoti, "Planning in a hierarchy of abstraction spaces", *AIJournal*, Vol.5, 1974, pp.115-135.

[Schank 77] Roger C. Schank et Robert P. Abelson, *Scripts, Plans, Goals and Understanding*, Lawrence Erlbaum, Hillsdale, N.J., 1977.

[Shafer 76] G. Shafer, *A Mathematical theory of evidence*, Princeton University Press, Princeton, 1976.

[Shafer 87] G. Shafer et R. Logan, "Implementing Dempster's rule for hierarchical Evidence", *Artificial Intelligence*, Vol.33, 1987, pp. 271-298.

[Shieber 85] Stuart M. Shieber, "An introduction to unification-based approaches to grammar", *Actes 23rd Annual Meeting of the A.C.L.*, Univ. of Chicago, 8 Juil., 1985.

[Shoham 87] Y. Shoham, "Temporal Logics in A.I. : Semantical and Ontological Considerations", *A.I. Journal*, Vol.33 n°1, 1987, pp.89-104.

[Shoham 88a] Y. Shoham et D. McDermott, "Problems in formal Temporal Reasoning", *A.I. Journal*, Vol.36, 1988, pp.49-61.

[Shoham 88b] Y. Shoham, "Chronological Ignorance: Experiments in NonMonotonic Temporal Reasoning", *AIJournal*, Vol.36, 1988, pp.279-331.

[Siroux 85] J. Siroux et D. Gillet, "A system for man-machine communication using speech", *Speech Communication*, Vol.4 n°4, 1985, pp. 289-315.

[Sowa 84] J.F. Sowa, *Conceptual Structures : Information Processing in Mind and Machine*, Addison Wesley 1984.

[Stalnaker 84] R. Stalnaker, "L'assertion", in : Frédéric Nef ed., *L'analyse logique des langues naturelles, 1968-1978*, CNRS, Paris, 1984.

[Touretzky 86] D.S. Touretzky, *The mathematics of inheritance systems*, Morgan Kaufman, Los Altos, 1986.

[Tsang 87] Edward P.K. Tsang, "Time structures for A.I.", *Proc. IJCAI 87*, 1987, pp.456-461.

[Tyler 86] L.K. Tyler et W. Marslen-Wilson, "The Effects of Context on the Recognition of Polymorphic Words", *Journal of Memory and Language*, Vol.25, 1986, pp.741-752.

[Véronis 88] Jean Véronis, *Contributions à l'erreur dans le dialogue Homme-Machine en langage naturel*, Thèse, Université d'Aix-Marseille III, 1988.

[Winograd 73] T. Winograd, "A procedural model of language understanding", in *Computer Models of thought and language*, R.C. Schank et K.M. Colby (eds), Freeman, San Francisco, 1973.

[Wolf 80] J.J. Wolf et W.A. Woods, "The HWIM Speech Understanding System", in : W. A. Lea, eds., *Trends in speech recognition*, Prentice Hall, 1980, pp. 316-339.

[Woods 70] W.A. Woods, "Transition Network Gramars for Natural Language Analysis", *Communication of the ACM*, Vol.13, pp.561-606, 1970.

[Yip 85] Kenneth Man-kam Yip, "Tense, aspect and the cognitive representation of time", *proc. IJCAI 85*, Los Angeles, pp.806-814.

[Zadeh 65] L.A. Zadeh, "Fuzzy sets", *Information and control*, Vol.8, 1965, pp. 338-353.

APPENDICE : TRACE DU GESTIONNAIRE D'HYPOTHESES.

The predicate Primitives.

Table 1 shows a description of the primitives used in our system, from several basic features, about which we give complementary explanations :

CTR (Control) : the predicate concerns an action or a state,that involves an explicit control.

DYN (Dynamic) : the predicate induce a change in a world configuration.

ANI (Animate) : the predicate involves or not an animate entity.

MVT (Movement) : the primitive implies that an entity is moved.

TRS (Transitive) : the predicate needs at least two arguments.

PRO (Production) : transfer of an entity from the controller of the action to another entity.

SPA (Space) : the primitive needs a spatial dimension expression.

Verb Type	Verb Primitive	CTR	DYN	ANI	MVT	TRS	PRO	SPA	Example
ACT	MVT1	+	+	+	+	-			to go
	MVT2	+	+	+	+	+			to bring
	ACTANIME	+	+	+	-	+			to vaccinate
	ACTOBJET	+	+	+	-	+			to sign
	ECHPROD	+	+	+	-	+	+		to give
	ECHOBT	+	+	+	-	+	-		to obtain
	ATRANS	+	+	+	-	-			to vote
STATE	LOCATION1	-	-		-	-		+	to find o.s.
	LOCATION2	+	-	+	-	+		+	to keep
	EXTENSION	+	-	+	-	+		-	to know
	STATUT1	-	-	+	-	-		-	to be french
	STATUT2	-	-	-	-	-			to be valid
	MESURE1	-	-	+	-	-		-	to be old
	MESURE2	-	-	-	-				to cost
PROCESS	PROCESS1	-	+	+	-	+			to change
	PROCESS2	-	+	-	-				to be destroyed

Table 1.

The primitives cases.

The principal cases involved in the definition of a primitive are described with the semantic restrictions that terms resulting from a syntactic parsing shall fulfill in order to be a plausible candidate for that case.(see table 2)

AGT (Agent) : in general, the controlling entity of a predicate.
semantic restriction : +Animate.

PAT (Patient) : the entity affected by a predicate.
semantic restriction : +Animate.

OBJ (Object) : the entity affected by a predicate.
semantic restriction : -Animate.

BEN (Benificiary) : entity to which something is transferred, or to which an action is applied.
semantic restriction : +Animate.

SRC (Source) : entity from which another entity is moved.
semantic restriction : +Location.

Verb Primitive	AGT	PAT	OBJ	BEN	SRC	DES
MVT1	+	-	-	-	+	+
MVT2	+	-	+	-	+	+
ACTANIME	+	-	-	+	-	-
ACTOBJET	+	-	+	o	-	-
ECHPROD	+	-	+	+	-	-
ECHOBT	o	-	+	+	-	-
ATRANS	+	-	-	o	-	-
LOCATION1	-	+/-	-/+	-	-	-
LOCATION2	-	+	+	o	-	-
EXTENSION	-	+	+	-	-	-
STATUT1	-	+	-	-	-	-
STATUT2	-	-	+	-	-	-
MESURE1	-	+	-	-	-	-
MESURE2	-	-	+	o	-	-
PROCESS1	-	+	o	-	-	-
PROCESS2	-	-	+	-	-	-

Table 2.

?	summary of commands.
init	initialization of scores as word frequencies.
univers	description of every node of the hierarchy.
lexique	description of the leaves (the lexems) of the hierarchy.
val_niv	minimum level under witch a word is not taken into account.
niveau	positionning of preceding level.

The scores are initialised from a corpus within the domain of our application, they are rounded off for readability, and have no exact linguistic signification, except for a pure relative comparison of words on a frequency base. "Number" indicates the number of occurences in the corpus used.

The hypotheses that select a sub-lexicon use the operators **non, union** and **intersection,** with some sub-lexicon or particular nodes of the hierarchy as argument. If the hypothesis is given with a score, it is considered as a contribution of information.

```
? (gestion)                          ;entering the hypotheses manager

GESTIONNAIRE D'HYPOTHESES

hypothese ? ?                        ; available commands
(init univers lexique mode val_niv niveau ?)

hypothese ? univers                                  ;description of the universe

;NODE          NUMBER   SCORE      ANCESTERS
                                                     ; the lexicon
document       5        .0123      anime- nom       ; non animate noun
je             249      .6118      anime+ pronom    ; pronoun
mesurer        2        .0049      mesure1 verbe    ; verb
compliquer     1        .0025      process2 verbe
important      1        .0025      statut2 adjectif ; the same primitive instanciates
importance     1        .0025      statut2 nom      ; in three possible grammatical forms
importer       4        .0098      statut2 verbe    ;
habitation     1        .0025      statut1 nom
habiter        5        .0123      statut1 verbe
appartenir     4        .0098      location1 verbe
remis          1        .0025      exchprod nom
remettre       2        .0049      exchprod verbe
remercier      21       .0516      actanime verbe
changement     1        .0025      process1 nom
changer        8        .0197      process1 verbe
couter         1        .0025      mesure2 verbe
savoir         29       .0712      extension verbe
```

```
garder      2       .0049     location2 verbe
vote        1       .0025     atrans nom
obtention   2       .0049     echobt nom
obtenir     24      .0590     echobt verbe
donner      21      .0516     exchprod verbe
signer      1       .0025     actobjet verbe
apporter    2       .0049     mvmt2 verbe
aller       18      .0442     mvmt1 verbe

pronom      249     .6118     cat_gram      ; each grammatical category refers
adjectif    1       .0025     cat_gram      ; to a common node : cat_gram
nom         12      .0295     cat_gram
verbe       145     .3563     cat_gram
cat_gram    407     1                       ; the top node of one of an arborescence
anime-      5       .0123     anime
anime+      249     .6118     anime
anime       254     .6241     trait_sem
trait_sem   254     .6241

process2    1       .0025     process ani- pat- obj+          ; "process" is an
process1    9       .0221     process ani+ trs+ pat+ ins-     ; intermediate node
mesure2     1       .0025     ctrtrs- ani- pat- obj+ tmp- mes+
            ; so is "ctrtrs-"
mesure1     2       .0049     ctrtrs- ani+ spa- pat+ obj- ben- +loc- tmp- mes+
statut2     6       .0147     ctrtrs- ani- pat- obj+ ben-
statut1     6       .0147     ctrtrs- ani+ spa- pat+ obj- ben- +loc- mes-
extension   29      .0712     ctrtrs spa- pat+ obj+ ben- loc- +mes-
location2   2       .0049     ctrtrs spa+ pat+ obj+ mes-
location1   4       .0098     ctrtrs- spa+ pat- pat- obj- obj+ +ben- loc+ ins- mes-
ctrtrs-     19      .0467     state ctr- trs-
ctrtrs      31      .0762     state ctr+ ani+ trs+
atrans      1       .0025     act mvt- trs- agt+ pat- obj- +src- des- ins- mes-
echobt      26      .0639     dep_g pro- pat- obj+ ben+ src- +des- ins- mes-
exchprod    24      .0590     dep_g pro+ agt+ pat- obj+ ben+ +src- des- ins- mes-
actobjet    1       .0025     dep_g agt+ pat- obj+ src- des- +mes-
actanime    21      .0516     dep_g agt+ pat- obj- ben+ src- +des-
mvmt2       2       .0049     mvt_g trs+ agt+ pat- obj+ ben- +src+ des+ loc- ins- mes-
mvmt1       18      .0442     mvt_g trs- agt+ pat- obj- ben- +src+ des+ loc- ins- mes-
dep_g       72      .1769     act mvt- trs+
mvt_g       20      .0491     act mvt+
process     10      .0246     ctr- dyn+ mvt- agt- ben- src- +des- mes- med- goal-
state       50      .1228     dyn- mvt- agt- src- des- med- +goal-
act         93      .2285     ctr+ dyn+ ani+
goal-       60      .1474     goal
goal        60.     .1474                   ; partial top of an arborescence
med-        60      .1474     med
med         60      .1474
mes-        123     .3022     mes
mes+        3       .0074     mes
mes         126     .3096
ins-        84      .2064     ins
ins         84      .2064
tmp-        3       .0074     tmp
tmp         3       .0074
```



```
loc-        57      .1400     loc
loc+        4       .0098     loc
loc         61      .1499
des-        133     .3268     des           ; each primitive feature is divided in
des+        20      .0491     des           ; two nodes : +feature/-feature ,
des         153     .3759                   ; indicating its presence or absence
src-        133     .3268     src
src+        20      .0491     src
src         153     .3759
ben-        77      .1892     ben
ben+        71      .1744     ben
ben         148     .3636
obj-        52      .1278     obj
obj+        96      .2359     obj
obj         148     .3636
pat-        109     .2678     pat
pat+        48      .1179     pat
pat         157     .3857
agt-        60      .1474     agt
agt+        67      .1646     agt
agt         127     .3120
spa-        37      .0909     spa
spa+        6       .0147     spa
spa         43      .1057
pro-        26      .0639     pro
pro+        24      .0590     pro
pro         50      .1228
trs-        38      .0934     trs
trs+        114     .2801     trs
trs         152     .3735
mvt-        133     .3268     mvt
mvt+        20      .0491     mvt
mvt         153     .3759
ani-        8       .0197     ani
ani+        141     .3464     ani
ani         149     .3661
dyn-        50      .1228     dyn
dyn+        103     .2531     dyn
dyn         153     .3759
ctr-        29      .0712     ctr
ctr+        124     .3047     ctr
ctr         153     .3759

hypothese ? lexique                 ;description of the lexicon

;NODE       NUMBER  SCORE     ANCESTERS

 document   5       .0123     anime- nom      ; "document"
je          249     .6118     anime+ pronom   ; "I"
mesurer     2       .0049     mesure1 verbe   ; "to measure"
compliquer  1       .0025     process2 verbe -                ; "to complicate"
important   1       .0025     statut2 adjectif ; "important"
importance  1       .0025     statut2 nom     ; "importance"
importer    4       .0098     statut2 verbe   ; "to matter"
```

```
habitation    1      .0025    statut1 nom      ; "living"
habiter       5      .0123    statut1 verbe    ; "to live"
appartenir    4      .0098    location1 verbe  ; "to belong"
remise        1      .0025    exchprod nom     ; "delivery"
remettre      2      .0049    exchprod verbe   ; "to deliver"
remercier     2      .0516    actanime verbe   ; "to thank"
changement    1      .0025    process1 nom     ; "change"
changer       8      .0197    process1 verbe   ; "to change"
couter        1      .0025    mesure2 verbe    ; "to cost"
savoir        29     .0712    extension verbe  ; "to know"
garder        2      .0049    location2 verbe  ; "to keep"
vote          1      .0025    atrans nom       ; "vote"
obtention     2      .0049    echobt nom       ; "obtaining"
obtenir       24     .0590    echobt verbe     ; "to obtain"
donner        21     .0516    exchprod verbe   ; "to give"
signer        1      .0025    actobjet verbe   ; "to sign"
apporter      2      .0049    mvmt2 verbe      ; "to bring"
aller         18     .0442    mvmt1 verbe      ; "to go"

                              ;some sub-lexicons .....

hypothese ?   ((union ctr- mvt+) ())
              ;words indicating a motion or the absence of control
aller         -> score = .0442      ; this word will be first searched on the signal
changer       -> score = .0197
habiter       -> score = .0123
appartenir    -> score = .0098
importer      -> score = .0098
mesurer       -> score = .0049
apporter      -> score = .0049
couter        -> score = .0025      ; these words will be searched at last
changement    -> score = .0025      ;
important     -> score = .0025      ;
compliquer    -> score = .0025      ;
habitation    -> score = .0025      ;
importance    -> score = .0025      ;

hypothese ?   ((inter (non verbe) (union ctr+ ben-)) ())
                    ;anyword except verbs expressing a controled operation
                    ; with a beneficiary case
obtention     -> score = .0049      ; very probable in our application
vote          -> score = .0025
changement    -> score = .0025
remis         -> score = .0025
habitation    -> score = .0025
importance    -> score = .0025
important     -> score = .0025

hypothese ?   niveau 0.02           ;setting the minimum selection level
hypothese ?   (verbe ())            ;verbs available
savoir        -> score = .0712
obtenir       -> score = .0590
donner        -> score = .0516
remercier     -> score = .0516
```



```
aller         -> score = .0442

hypothese ?   ((inter mvt+ ctr+) .9)
              ;contribution of information for controlled motions

hypothese ?   (verbe ())
              ;new distribution for verbs after hypothesis
aller         -> score = .3066      ; "aller" has been brought forward by the hypothesis
savoir        -> score = .0494
obtenir       -> score = .0409
remercier     -> score = .0358
donner        -> score = .0358
apporter      -> score = .0341

hypothese ?   ( )           ;end of session
```

NOM DE L'ETUDIANT : ROMARY Laurent

NATURE DE LA THESE : Doctorat de l'Université de NANCY I en Informatique

VU, APPROUVE ET PERMIS D'IMPRIMER

NANCY, le 10 AVR. 1989 n° 613

LE PRESIDENT DE L'UNIVERSITE DE NANCY I

M. BOULANGÉ

Résumé.

Ce rapport présente un ensemble de recherches dont l'objectif est de mettre en œuvre des systèmes de dialogue homme-machine possédant un comportement aussi naturel que possible. La réflexion proposée s'articule autour de trois axes principaux :

- dans un premier temps, nous montrons, sur la base d'une étude particulière, que la multiplicité des modèles de représentation utilisés dans les systèmes de dialogue actuels conduit à une rigidité excessive de fonctionnement pour ceux-ci.

- nous cherchons, dans le cadre des sciences cognitives, s'il existe un modèle unifié applicable à l'ensemble des informations linguistiques et déductives manipulées dans un dialogue. Les difficultés que nous rencontrons alors nous obligent à nous limiter à un domaine particulier : l'étude des informations temporelles, qui touchent tous les niveaux de traitement d'un système de dialogue.

- nous proposons un modèle cognitif de représentation du temps qui permet effectivement de considérer sous un même angle des phénomènes liés au langage, à la représentation des connaissances et au raisonnement.

Le modèle proposé semble être une base possible pour une étude plus générale du dialogue homme-machine, en marge des approches modulaires classiquement envisagées.

Mots-clés : dialogue homme-machine, cognition, langage, temps.