

# Génération d'actes illocutoires pour le dialogue

Jean Caelen, Ludovic Imberdis

Laboratoire CLIPS - IMAG  
Domaine universitaire, BP 53  
38041 Grenoble cedex 9  
Courriel : Jean.Caelen@imag.fr

---

## Résumé

Les sorties parlées d'un système de dialogue homme-machine sont souvent d'une qualité médiocre pour deux raisons : le texte généré par le générateur ne tient souvent pas compte de l'utilisateur et la synthèse de la parole à partir de ce texte, donne souvent une prosodie de lecture. L'article décrit les principes d'introduction d'une composante pragmatique en amont du système de génération afin de tenir compte des degrés de force et des buts illocutoires. En effet, les études en génération portent généralement sur les réalisations linguistiques du contenu propositionnel de la réponse (c'est-à-dire le sujet du message) mais très rarement sur la constitution de l'acte énonciatif qui est généralement abandonnée au moteur de dialogue de l'application. Nous développons cet aspect au cœur même de la génération afin de fournir un outil performant et surtout pour prendre en compte les aspects pragmatiques de manière précoce dans le dialogue. La conclusion est que notre système, grâce à son module de traitement pragmatique, a la capacité de générer des énoncés plus naturels. La génération est basée non plus sur le contenu informatif seul, mais aussi sur une situation pragmatique, exprimée dans un cadre bien défini. Notre approche montre l'intérêt de prendre en compte les paramètres illocutoires dans le cadre du dialogue.

---

## 1. Introduction

La technique *text-to-speech* pour la synthèse de la parole consiste à générer un énoncé oral à partir d'un texte. Les systèmes utilisant cette technique obtiennent souvent, une prosodie de type lecture. L'approche, qui consisterait simplement à placer un tel système *text-to-speech* en sortie d'un module de dialogue, serait doublement préjudiciable dans une situation de dialogue homme-machine car (a) les interventions des interlocuteurs ne sont pas des textes et (b) leurs répliques ne sont pas des lectures. En effet on ne dialogue pas comme on lit, on ne dialogue pas non plus comme on écrit : les formes linguistiques sont différentes (au niveau syntaxique et au niveau prosodique, mais aussi par un emploi plus fréquent d'ellipses, de formes emphatiques, etc.), et surtout ces formes véhiculent une composante perlocutoire évidente. Plus que la recherche du *naturel* qui n'est peut-être pas une fin en soi, il faut viser l'intelligibilité et l'adéquation ou la justesse d'une réponse qui procède de sa pertinence ; on sait également qu'une réponse peu pertinente n'est pas compréhensible. Au niveau du dialogue homme-machine il y a donc nécessité de contrôler cette pertinence pour deux raisons : (a) pour diminuer le coût cognitif de l'utilisateur (avec des retombées évidentes sur la performance d'exécution de la tâche, la diminution de la fatigue, etc.) et, (b) pour respecter le principe ergonomique d'*observabilité* qui stipule que l'utilisateur doit avoir une conscience claire des véritables capacités de la machine.

La solution que nous proposons passe par deux étapes : la première, qui précède la synthèse proprement dite est d'adapter un système de génération effectivement au dialogue et non plus simplement à la réalisation linguistique d'un contenu, et la seconde est de faire intervenir les niveaux pragmatiques dans la génération de la prosodie. Cette approche qui fait donc appel à la pragmatique met l'accent sur l'aspect primordial de l'intersubjectivité que négligent la plupart des systèmes pour l'instant. - les éléments pragmatiques que sont seulement capables de gérer les interfaces homme-machine se situant ou bien à un niveau de planification, Pollack M. (1990), par exemple, ou bien au niveau des problèmes de résolution des références.

Notre travail s'inscrit dans une perspective de dialogue oral pour la génération d'énoncés dans laquelle nous tentons de concilier la génération de texte classique proprement dite (sémantique, syntaxe) et la pragmatique illocutoire, basé sur les théories d'Austin (Austin, 1962). Nos travaux reposent en partie sur le noyau de génération développé par l'équipe Cristal-Gresec de l'université Stendhal de Grenoble, (Balicco L. 1993 et Ponton C. 1996), sur lequel nous greffons un module de traitement pragmatique illocutoire (désormais appelé *module pragmatique* dans la suite du texte).

## 2. La pragmatique illocutoire dans le contexte de la génération

Austin (Cf. Austin, 1962) fut l'un des premiers à remettre en cause la conviction selon laquelle le langage a pour fonction de décrire un état de faits et doit donc conséquemment produire des affirmations susceptibles d'être vraies ou fausses (Cf. Searle, 1969). Austin distingue dans l'action qui consiste à produire un énoncé, 3 types d'actes de langage : l'acte locutoire, l'acte illocutoire et l'acte perlocutoire, chacun de ces actes représentant respectivement le contenu de l'acte, ses visées et moyens, et ses effets. Nous retenons d'Austin et de Vanderveken (Vanderveken, 1988) qui a formalisé la logique illocutoire, que chaque énoncé est, à la base, formé de trois constituants principaux : **un contenu propositionnel**, un **but** (que le locuteur poursuit) et le **degré de force** qu'il utilise pour réaliser ce but. Le contenu propositionnel est du niveau du locutoire, le but et la force du niveau illocutoire. Malgré ses qualités, cette approche est essentiellement monologique, c'est-à-dire qu'elle ne prend pas en compte le locuteur dans la dynamique du dialogue. Nous devons donc l'étendre dans un cadre plus général prenant en compte les autres éléments du dialogue qui permettent d'aborder le problème de l'intersubjectivité.

La figure 1 présente le cadre que nous avons choisi de donner au dialogue et la relation qui existe entre chacun de ses éléments avec les théories d'Austin.

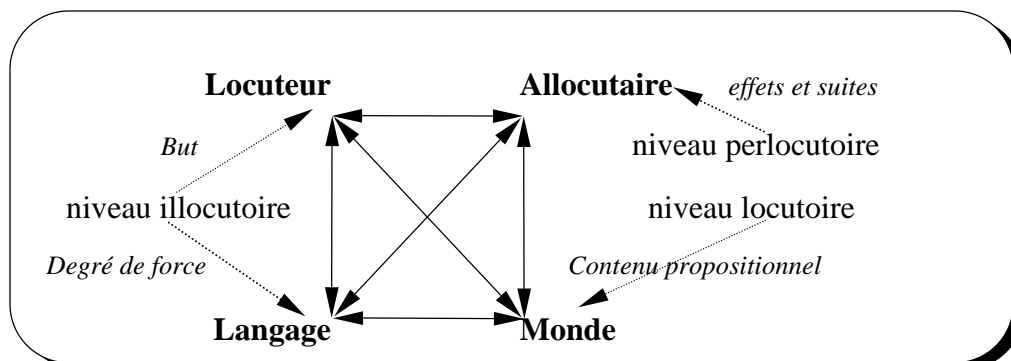


Figure 1. Cadre pour le dialogue et liaison avec les théories d'Austin.

Ce cadre représente les éléments qui entrent en jeu à chaque tour de parole. C'est à dire aussi bien les relations qui interviennent entre le locuteur et l'allocutaire - et c'est le point sur lequel nous mettons l'accent - que leur connaissance respective du langage et que les rapports

qu'ils peuvent avoir au monde. Deux pôles se dégagent à partir de ce cadre et dans la perspective du dialogue :

- d'un côté, les aspects purement langagiers et référentiels - représenté par le *langage* et le *monde* - que Searle (Cf. Searle, 1969) a mis en évidence par la notion de direction d'ajustement entre les mots et le monde et qui sont déjà pris en compte au niveau des générateurs (au moins pour l'aspect référentiel),

- de l'autre côté, les aspects intersubjectifs - représentés par le *locuteur* et l'*allocutaire* - qu'il est indispensable de faire intervenir dans une situation de dialogue au cours de laquelle, ils coordonnent leurs actions (Vernant, 1992) et règlent leur jeu inter-social.

Selon l'hypothèse de Searle « tout ce que l'on veut dire peut être dit » (principe d'*exprimabilité*) nous postulons que chacun des éléments du cadre ainsi que les relations qu'ils entretiennent peuvent se projeter dans la langue au niveau de l'énoncé. Inversement, les énoncés véhiculent ces éléments que l'allocutaire saura reconnaître et interpréter. Cette approche résolument linguistique s'inscrit dans la perspective du marquage illocutoire et interactif des énoncés des travaux de Roulet E. (1980), Rubattel C. (1980) et De Spengler N. (1980).

### **3. Réalisation**

La visée globale de l'utilisation de la pragmatique illocutoire en dialogue homme-machine est d'améliorer la pertinence des répliques de la machine, en agissant sur la forme et sur le contenu des énoncés générés. Plus prosaïquement, notre réalisation s'attache à mettre en valeur les structures linguistiques nécessaires à la prise en compte d'une formalisation des actes de langage, dans un cadre restreint au dialogue. L'architecture générale de notre système est donnée sur la figure 2.

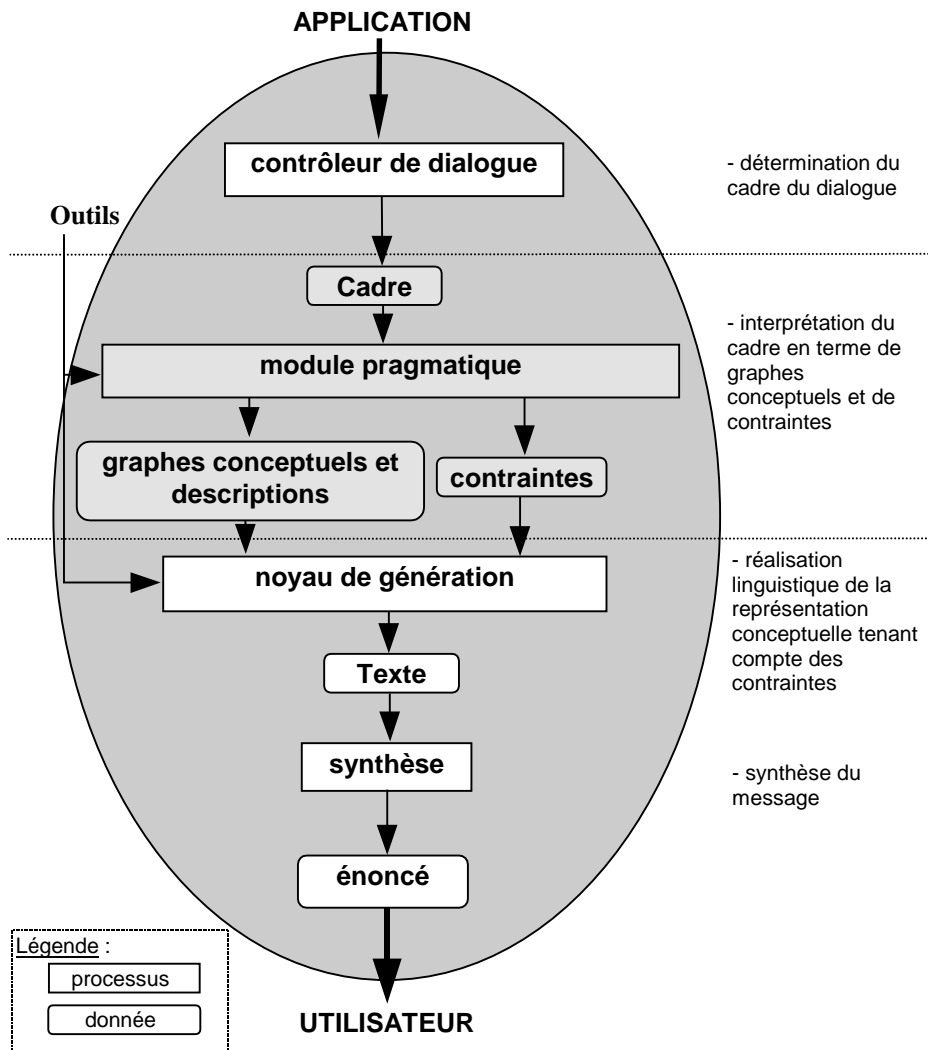


Figure 2. architecture générale du système.

Formellement nous appelons  $F$  la forme et  $P$  le contenu propositionnel d'un énoncé. Il s'agit pour la machine (jouant le rôle du locuteur  $L$ ), de produire l'énoncé  $F(P)$  dans le contexte du dialogue, c'est-à-dire sous la contrainte de réalisation du but et à l'intention de l'allocutaire  $A$ . La forme (ou force illocutoire chez Vanderveken) est définie par le but illocutoire  $b$  et le degré de force  $f$ . On pose  $F = \langle b, f \rangle$ . Nous supposons que  $F$  et  $P$  sont indépendants ce qui permet de réaliser le système de génération à l'aide de modules indépendants. Ainsi on peut envisager quatre étapes dans le processus :

- la génération du contenu propositionnel  $P$  en fonction de l'avancée du dialogue et de la tâche,
- le calcul du but illocutoire  $b$ , en fonction du but de la tâche,
- le calcul du degré de force  $f$ ,
- la prise en compte des rôles que jouent  $A$  et  $L$  et de leur relation  $Rel(A,L)$ .

Une fois ces informations obtenues, on procède à la mise en forme du message de sortie, c'est-à-dire à la production proprement dite de l'énoncé en ajustant sa forme finale à la force illocutoire.

Nous appelons cadre illocutoire la structure  $F(P, A, L, Rel)$  où  $F = \langle b, f \rangle$ .

Nous avons défini une syntaxe sur ce cadre de façon à ce que le *module pragmatique* puisse en traiter l'ensemble des éléments. Il prend en entrée les éléments du cadre et fournit en sortie tous les éléments nécessaires au *noyau de génération* pour générer l'énoncé proprement dit en langage naturel.

## 4. Représentation et prise en compte des éléments du cadre dans le système

### 4.1. Représentation

Nous nous fondons sur les représentations utilisées par Ponton C. (1996) dans son *noyau de génération*. Il s'agit des graphes conceptuels (Cf. Sowa J. 1984) dont le modèle trouve ses fondements en linguistique et en psychologie. C'est pourquoi il est particulièrement bien adapté au traitement du langage naturel. Nous ne discuterons pas ce point dans cet article.

Il existe par ailleurs au sein du noyau de génération un module de traduction de formalisme qui permet de traiter les CMR - Common Meaning Representation - et les descriptions fonctionnelles. Au niveau du noyau de génération, les graphes conceptuels sont représentés en Prolog sous la forme d'un prédicat à deux arguments : une liste de variables de *description* et la définition même du *graphe*. Nous précisons ce point qui a son importance dans le fonctionnement du module pragmatique. En effet, les descriptions associées à chaque graphe conceptuel regroupent l'ensemble des caractéristiques de la proposition associée au graphe conceptuel. Elles permettent par exemple de préciser la forme (assertive, interrogative ou impérative) de la proposition et le temps verbal à employer. C'est aussi cette approche des graphes conceptuels qui a influé sur le choix du noyau de génération comme base de développement. En effet ces descriptions sont une première approche des caractéristiques linguistiques globales et essentielles de l'énoncé qui permettent une maîtrise des effets. Les descriptions vont donc être un des points de liaison entre le générateur proprement dit et le module pragmatique.

### 4.2. Le locuteur et l'allocutaire

La notion d'intersubjectivité dont Benveniste E. a montré que la condition était le langage, est prise en compte dans les rôles respectifs que jouent L et A, ainsi que dans leur relation Rel. Ceci implique et justifie qu'il soit possible de marquer linguistiquement ces rôles et cette relation. C'est en parlant, qu'un individu se pose comme un « je » devant un « tu ». « Nous parlons à d'autres qui parlent, telle est la réalité humaine » commente Benveniste E. (in Benveniste, 1966). Qui plus est, du point de vue de l'action, Vernant D. considère que cette dernière n'existe à travers l'acte de langage, que dans l'interprétation qui en est faite par celui qui fait l'action.

Dans cette ligne de pensée, nous avons donc recensé l'ensemble des rôles que peuvent tenir L et A pour les prendre en compte au niveau de la génération. Selon les rôles respectifs de L et A (apparaissant dans la relation Rel(L,A)), différentes formes linguistiques peuvent marquer la subjectivité. Au niveau interne le module pragmatique est capable de traiter aussi bien les formes personnelles directes, représentatives des rôles (c'est à dire « je » et « tu ») que les formes conceptuelles plus traditionnelles, même si la subjectivité n'est réellement marquée qu'à travers les formes personnelles - nonobstant les remarques de Moeschler J. & Reboul A. (1994) quant à la subjectivité de la troisième personne.

### 4.3. La forme

#### a) la force

Au niveau de la force, nous utilisons une échelle numérique allant de 0 à 5, ainsi que des

indicateurs de dynamicité<sup>1</sup>. Ces indicateurs permettent de préciser si la force est croissante, décroissante ou stable, d'une intervention à la suivante. La prise en compte de ces phénomènes nécessite la gestion de l'historique du dialogue. Etant donné que les réalisations linguistiques de la force sont fortement dépendantes des relations qui unissent le locuteur et l'allocutaire, cela impose aussi de disposer d'une représentation de l'utilisateur dans la machine (mais pour l'instant, nous n'avons pas encore de modèle de l'utilisateur sophistiqué dans notre système).

#### b) le but illocutoire

Un but se définit vis-à-vis de la tâche sur laquelle porte le dialogue. On appelle *but* un état du monde que désire atteindre le locuteur, aussi bien qu'un état mental auquel il souhaite arriver (par exemple connaître une information, acquérir un savoir-faire). Vis-à-vis d'un acte particulier, accompli dans le but de réaliser une tâche, on appelle *but illocutoire* la visée actionnelle ou communicationnelle du locuteur à l'adresse de l'allocutaire.

Le passage du *but* au *but illocutoire* permet de s'affranchir des contraintes qui existent au niveau de chaque application puisque des buts particuliers nous passons aux classes génériques bien définies des buts illocutoires. A charge pour l'instant au moteur de dialogue, d'opérer la transition entre les buts - par exemple, de réservation, de renseignement, etc. dans une application de réservation hôtelière - aux buts illocutoires qui permettent l'aboutissement à de tels buts.

Dans ces conditions, en réduisant notre propos au dialogue homme-machine opératif, les différents buts illocutoires que nous avons retenus se répartissent en trois grandes classes :

- le faire-faire, qui correspond à une demande de réalisation d'action par l'allocutaire,
- le faire-savoir et le faire-croire, qui correspondent à une réponse à l'allocutaire ou à une assertion de la part du locuteur – c'est un apport d'informations certaines ou incertaines (ou conjecturales),
- le faire-faire-savoir et le faire-faire-croire, qui correspond à une demande d'information (mais aussi de confirmation).

Ainsi, par exemple, à partir de  $P = \boxed{\text{fermer}} \Rightarrow \text{objet} \Rightarrow \boxed{\text{porte}}$  on obtient les énoncés suivants :

- avec but illocutoire = faire-faire            « Ferme la porte ! »
- avec but illocutoire = faire-savoir        « La porte est fermée. »
- avec but illocutoire = faire-faire-savoir   « Est-ce que la porte est fermée ? »

## 5. Module de traitement pragmatique illocutoire

Les traitements qu'effectue ce module permettent d'obtenir, à partir des éléments du cadre du dialogue, une représentation conceptuelle interprétable par le noyau de génération. Il s'agit donc de traduire tous les éléments du cadre en termes de représentations conceptuelles puis de produire une sortie pour le noyau de génération. Pour cela le module pragmatique travaille à l'aide de quatre opérateurs de base. La fonction de ces opérateurs de base est de transformer la représentation conceptuelle de départ qui fait partie des éléments du cadre en une

---

<sup>1</sup> On pourra trouver dans la littérature maintes autres échelles de force, pour ne citer par exemple que (Boissier O. 1993) qui utilise les valeurs {null, informing, bargaining, persuading, commanding}. Le manque de consensus dans la prise en compte de la force prouve la difficulté à définir une échelle stable.

représentation plus complexe dans laquelle interviennent les autres éléments du cadre, c'est à dire à la fois la prise en compte du but, de la force et la représentation du locuteur et de l'allocutaire. Ces opérateurs de base sont :

- *l'adjonction* : l'information pragmatique se rajoute au contenu propositionnel proprement dit sous la forme d'une proposition complémentaire : « **je te dis que** "la station est sur la table" », « **je veux que tu** "fermes la porte" », etc. C'est essentiellement la prise en compte du but qui fait intervenir ce type de phénomènes ;

- *la substitution* : l'information pragmatique passe par le remplacement d'un composant par un autre ; par exemple la pronominalisation (ou son inverse) : de « **je** suis bien malheureux... » à « **ton pauvre père** est bien malheureux... » ;

- *la suppression* : l'information pragmatique utilise les possibilités de raccourci du langage (notion d'ellipse notamment) pour atteindre les effets voulus : « Le sel ! » devant être compris comme « **Passe-moi** le sel ! », puis « le sel, sur la table ! » signifiant « **Remets** le sel sur la table ! », etc. ;

- *la modification* : l'information pragmatique peut obliger à des transformations plus radicales de l'énoncé avec par exemple les phénomènes d'indirection tel que « Avez-vous du sel ? ».

Le fonctionnement général du module pragmatique est le suivant :

a) Le module commence par interpréter le but illocutoire *b* pour compléter le graphe conceptuel qui représente *P* à exprimer en lui adjoignant un certain nombre de caractéristiques qui sont du niveau des descriptions. Ainsi en règle générale la différence entre un faire-savoir et un faire-faire-savoir pour le même *P* va se faire sur la forme linguistique de la phrase : interrogative ou non.

b) L'étape suivante consiste à prendre en compte le *degré de force* pour calculer les contraintes, utiliser les opérateurs de base possible et exprimer le degré de force au niveau des concepts.

c) En fonction de *L*, *A* et *Rel(L,A)* des branches supplémentaires peuvent être greffées au graphe conceptuel initial *P* pour pouvoir exprimer linguistiquement la subjectivité dans l'énoncé. Ensuite, et selon leur relation, *L* et *A* sont pris en considération.

Le module s'appuie sur les capacités linguistiques du noyau de génération pour pouvoir *exprimer* l'ensemble des éléments en présence, en terme de forme de surface. Ceci est possible notamment par le biais des outils que partagent le module et le noyau, outils qui sont soit génériques soit spécifiques à une application, notamment au niveau du lexique. L'avantage du retour-arrière (backtracking de Prolog) est de permettre la résolution des incompatibilités qui peuvent apparaître entre des opérations de niveaux différents (par exemple, un choix effectué au niveau pragmatique peut entraîner l'échec de la génération mais uniquement au niveau de la réalisation linguistique) par une «remontée» dans l'arbre des solutions.

Pour clarifier l'exposé, nous donnons l'exemple d'un cadre de dialogue et d'une réalisation linguistique obtenue après traitement.

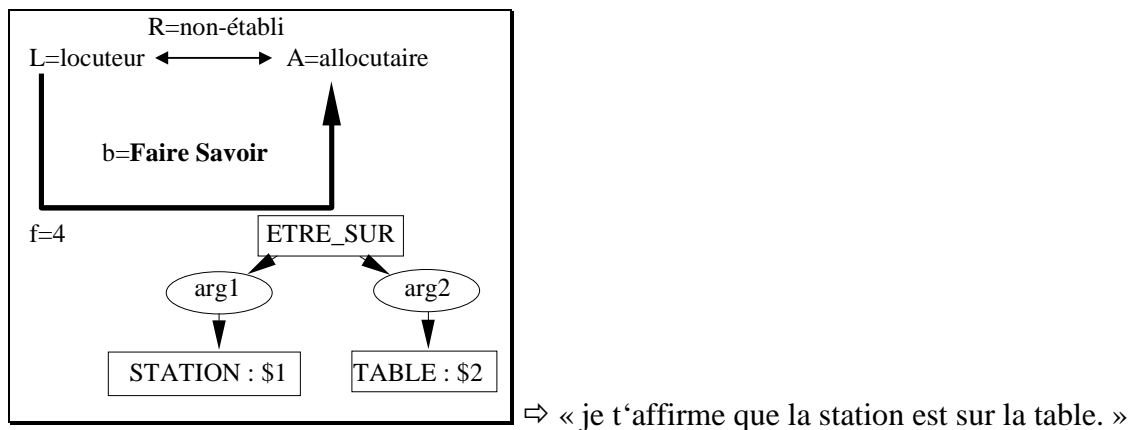


Figure 3. Exemple d'un cadre de dialogue.

## 6. Conclusion

Notre système, grâce à son module de traitement pragmatique, a la capacité de générer des énoncés plus naturels dans le dialogue. La génération est basée non plus sur le contenu informatif seul, mais aussi sur une situation pragmatique, exprimée dans un cadre bien défini. Notre approche montre l'intérêt de prendre en compte les paramètres illocutoires dans le cadre du dialogue. Outre les phénomènes que nous avons répertoriés et que notre système est capable de traiter, il existe une grande quantité de processus rhétoriques au sein du langage, qui restent encore à introduire. Nous sommes parfaitement conscients des limites de notre système face à la richesse et à la complexité du langage naturel (cependant tous les phénomènes n'ont certainement pas besoin d'être implémentés en dialogue homme-machine). D'autres phénomènes sont quant à eux, irrémédiablement trop complexes pour être traités dans le cadre limité que nous nous sommes fixés.

L'indépendance du module pragmatique permet de mettre en relief les phénomènes linguistiques dont la pragmatique fait usage, mais aussi d'avoir un outil facile à utiliser et à faire évoluer. Le fait d'obtenir un module de traitement pragmatique indépendant qui préexiste à la génération démontre que la pragmatique (au moins dans sa partie illocutoire) peut être un élément quasi indépendant qui améliore la qualité des réponses de la machine face aux demandes de l'utilisateur mais aussi au regard des exigences du concepteur de l'application. Le contenu, dans cette perspective n'est plus désormais le point central mais devient un élément d'un schéma plus global et plus riche.

## 7. Bibliographie

AUSTIN, John (1962), *How to do things with words*, Oxford, Oxford University Press.

BALICCO Laurence (1993), *Génération de répliques en français dans une interface homme-machine en langue naturelle*, thèse de troisième cycle, Université Pierre Mendès France, Grenoble.

BENVENISTE Emile (1966), *Problèmes de linguistique générale*, Paris, Gallimard.

BOISSIER Olivier (1993), *Problème du contrôle dans un système intégré de vision*, thèse de doctorat, I. N. P., Grenoble.

DE SPENGLER Nina (1980), "Première approche des marqueurs d'interactivité", *Actes de langage et structure de la conversation*, Unité de Linguistique française, Genève.

GARDINER A.H. (1989), *Langages et actes de langage. Aux sources de la pragmatique*,



Lille, Presse Universitaire de Lille.

GRAU Brigitte; SABAH Gérard; VILNAT Anne (1994), "Pragmatique et dialogue homme-machine", *TSI, numéro spécial sur les interfaces homme-machine*, pp. 9-30.

IMBERDIS Ludovic; CAELEN Jean (1996), "Actes de langage et dialogue Homme-Machine", *actes IHM'96*, Grenoble.

MOESCHLER Jacques; REBOUL A. (1994), *Dictionnaire encyclopédique de pragmatique*, Paris, Seuil.

POLLACK Martha (1990), "Plans as complex mental attitudes", *Intentions in communication*, MIT, Bradford books.

PONTON Claude (1996), *Formalisme et architecture d'un système ouvert multi-applications*, thèse de doctorat, Université Stendhal, Grenoble.

ROULET Eddy (1980), "Stratégies d'interaction et marqueurs illocutoires", *Actes de langage et structure de la conversation*, Unité de Linguistique française, Genève.

RUBATTEL Christian (1980), "Une analyse sémantique des verbes performatifs français", *Actes de langage et structure de la conversation*, Unité de Linguistique française, Genève.

SEARLE John (1969), *Speech acts*, Cambridge, Cambridge University Press.

SOWA John F. (1984), *Conceptual structures : information processing in man and machine*, Addison Wesley, Reading Mass.

VANDERVEKEN Daniel (1988), *Les actes de discours*, Bruxelles, Mardaga.

VERNANT Denis (1992), "Du dialogue", *recherches sur la philosophie et le langage n°14*. Paris, Vrin.