

# Traitement des incompréhensions et des malentendus en dialogue homme-machine

Jean Caelen<sup>1</sup>, Hoá Nguyen<sup>2</sup>

<sup>1</sup> Institut d'Informatique et Mathématiques Appliquées de Grenoble, CLIPS  
jean.caelen@imag.fr

<sup>2</sup> Vietnam National University, Hanoi  
hoa.nguyen@vnu.edu.vn

## Résumé

Traiter les erreurs en dialogue homme-machine est un problème difficile compte-tenu des multiples sources possibles depuis la reconnaissance de la parole jusqu'à la génération en passant par d'autres modules comme l'analyse sémantique, l'interprétation pragmatique ou la gestion du dialogue. Dans cet article, ce problème est envisagé dans le but d'apporter de la généralité et de la robustesse au système ; il est traité au niveau du contrôleur de dialogue. Les différents types d'erreurs sont d'abord identifiés et regroupés en deux catégories qui seules ont un sens vis-à-vis de l'utilisateur : les incompréhensions et les malentendus. Puis, ces deux catégories d'erreur sont traitées de manière spécifique pour que le système puisse générer une réponse convenable et intelligente à l'utilisateur, sans rupture de dialogue. L'expérimentation effectuée en appliquant cette approche au système de dialogue Mélina présente des résultats prometteurs pour traiter les erreurs en dialogue.

**Mots-clés** : incompréhension, malentendu, acte de dialogue, but de dialogue, stratégie de dialogue, dialogue homme-machine.

## Abstract

Miscommunication causes inconvenience within a spoken dialogue system. In this paper, we investigate this problem with regard to achieving genericity and robustness within dialogue management. Therefore, different error types occurring in a spoken dialogue system were first identified. Two markers were then proposed to detect and to trigger misunderstanding to the dialogue manager that should then choose an adequate strategy so as to generate an intelligent and suitable answer to the user. Moreover, several issues are investigated in order to process misinterpretations in the dialogue manager. An experiment performed using our approach to the spoken dialogue system Mélina shows promising results in processing miscommunication and proves the value of our approach.

**Keywords**: misunderstanding, misinterpretation, dialogue act, dialogue goal, dialogue strategy, human-computer dialogue.

## 1. Introduction

Dans la plupart des systèmes de dialogue homme-machine (DHM), l'un des problèmes majeurs concerne la robustesse, c'est-à-dire le maintien de l'interaction en dépit des erreurs de compréhension ou d'interprétation par le système : il s'agit de continuer le dialogue de manière constructive malgré les incompréhensions et de lever les malentendus entre le système et l'utilisateur (Bernsen *et al.*, 1998). Il existe de nombreux phénomènes qui peuvent provoquer des incidents au cours d'un dialogue, auxquels s'ajoutent les erreurs de traitement pouvant survenir dans tous les modules du système de dialogue homme-machine (SDHM) tels que la reconnaissance de la parole, le décodage sémantique de l'énoncé, l'interprétation pragmatique et quelquefois – au sein même du contrôleur de dialogue (CD) – les erreurs de

stratégie ou de but. Dans le système Mélina (Nguyen, 2005), qui a une architecture modulaire, les modules interagissent pour construire une solution de manière coopérative entre eux, chaque module ayant cependant un rôle spécifique et complémentaire : il n'y a donc pas de redondance dans le système. Pour traiter les erreurs, l'idée est donc de les faire remonter jusqu'au niveau du CD et de tenter d'y remédier à un moment opportuniste, en ayant une vue globale sur l'ensemble du dialogue. Cette idée s'inspire du comportement humain, dans lequel il semble que les erreurs soient détectées assez vite sous forme de contradictions mais ne sont pas corrigées tout de suite : le locuteur attend généralement la suite du discours pour faire des inférences manquantes, des vérifications ou rectifier des données mal comprises. Ce n'est qu'en dernier recours et devant une contradiction patente qu'il tente d'éclaircir la situation par un sous-dialogue *ad hoc* avec son interlocuteur.

La reconnaissance automatique de la parole, la compréhension et l'interprétation sont les modules les plus fragiles vis-à-vis des erreurs pour le DHM. À titre d'exemple, le dialogue illustré dans l'exemple 1 ci-dessous montre des erreurs provenant du module de reconnaissance de la parole (l'énoncé en gras et en italique est le texte obtenu par ce module pour des énoncés d'un utilisateur U. Le texte entre crochets est l'énoncé réel prononcé par l'utilisateur). Diverses erreurs bénignes se produisent ici (insertion délétion, substitution de mots, qu'il est possible de rectifier en tenant compte du contexte et de l'avancée de la tâche).

S : Bonjour, ici le service de réservation, je suis Mélina et vous ?
U : <b>bonjour jean</b> <bonjour ici jean>
S : Désolé, j'ai plusieurs personnes ayant votre prénom ! Quel est votre nom ?
U : <b>ce jugement caelen</b> <je suis Jean Caelen>
S : Bonjour monsieur le directeur ! Souhaitez-vous réserver ?
U : <b>il faut une salle de vingt personnes pour lundi prochain</b> <il me faut une salle de vingt personnes pour lundi prochain>
S : La salle aquarium pourrait convenir. A quelle heure s'il vous plaît ?
U : <b>quatorze heures</b> <à quatorze heures>
S : Je vous réserve donc la salle aquarium à 14 heures ce lundi 27 septembre
U : <b>très bien merci bien au revoir</b>
S : Au revoir monsieur Caelen à bientôt.

*Exemple 1. Un dialogue avec des erreurs non bloquantes*

Dans cet exemple, le système de dialogue peut prendre en compte les mots porteurs de sens et les erreurs peuvent être compensées par le module d'interprétation pragmatique. Ces cas concernent les erreurs les plus faciles à contourner et elles ne sont pas vraiment bloquantes pour le dialogue. Nous nous intéressons dans la suite de l'article aux cas les plus difficiles qui créent des incompréhensions ou des malentendus. Remarquons que la plupart des systèmes sont soumis à ces problèmes, par exemple, l'analyse du corpus du système CMU Communicator, réalisée par (Carpenter *et al.*, 2001), indique que 32 % des énoncés contiennent des erreurs de compréhension. D'autres résultats, sur d'autres systèmes, sont présentés dans le tableau 1 ci-dessous. Dans ce tableau, nous pouvons constater que 20 à 40 % des énoncés ne sont pas compris correctement par le système de dialogue.

Système de DHM	Ratio des erreurs sémantiques
Jupiter (Hazen <i>et al.</i> , 2002)	28 %
CMU Communicator (Carpenter <i>et al.</i> , 2001)	32 %
How May I Help You (Walker <i>et al.</i> , 2000)	36 %
CU Communicator (Rudnický <i>et al.</i> , 1999)	27 %
SpeechActs (Yankelovich <i>et al.</i> , 1995)	25 %

*Tableau 1. Proportion des énoncés contenant des incompréhensions dans divers systèmes*

## 2. Facteurs conduisant à un mauvais dialogue

Bien qu'intéressante pour le diagnostic des performances d'un système et son amélioration subséquente, il ne nous semble pas que cette typologie détaillée des facteurs défailants d'un SDHM soit une bonne piste de recherche pour augmenter la robustesse : elles seraient trop nombreuses à spécifier précisément et l'on ne serait pas à l'abri d'une mauvaise attribution d'erreur qui produirait une nouvelle erreur. Nous regroupons plutôt les erreurs d'un système de dialogue oral en deux catégories principales selon leurs conséquences sur le dialogue : l'incompréhension et le malentendu. En effet, la typologie précédente est peut-être pertinente pour le concepteur du système mais elle ne l'est pas vis-à-vis de l'utilisateur qui n'a pas à entrer dans la compréhension du système lui-même, il désire seulement poursuivre un dialogue de la manière la plus cohérente possible. Dans la communication quotidienne entre deux êtres humains, la conversation semble bien se dérouler jusqu'à ce qu'une contradiction éventuelle apparaisse. Cette contradiction est dû au fait que l'un parle de choses que l'autre ne comprend pas (incompréhension) dans le contexte de la conversation, ou bien qu'il parle d'une chose que l'autre interprète en une autre (malentendu). À un moment donné on constate donc une contradiction entre les interlocuteurs, ce qui est la marque de l'incompréhension ou du malentendu, qu'il s'agit de réparer sans remonter nécessairement aux causes profondes dans le système.

Ce sont seulement ces deux phénomènes que nous souhaitons détecter et corriger au niveau du contrôleur de dialogue. Certes, il peut y avoir des défauts dans le SDHM que l'on peut améliorer comme la couverture linguistique, la précision de la description ontologique, les algorithmes d'analyse, etc. pour rendre le SDHM plus robuste, mais ce n'est pas notre propos ici.

Définitions :

- **Incompréhension** : Formellement, l'incompréhension vient du fait que le système de DHM n'arrive pas à obtenir les schémas sémantiques ou pragmatiques Fp de manière correcte quand l'utilisateur prononce un énoncé. Une incompréhension se manifeste donc au niveau du CD par un acte Fp invalide, partiel ou vide.
- **Malentendu** : Formellement un malentendu se produit lorsqu'un énoncé est compris de manière différente par rapport à son intention. Les malentendus ne peuvent donc être traités que tardivement dans le décours du dialogue lorsqu'un faisceau de présomptions conduit à des contradictions évidentes par rapport aux faits ou lorsque l'utilisateur conteste certains actes. En fait l'acte Fp arrivant au CD est cohérent (valide en apparence) mais il produit des contradictions dans la base de faits du système.

## 3. Stratégie de traitement d'erreurs

Dans le cadre de cet article, nous ne nous intéressons qu'à la recherche d'une solution pour traiter les erreurs de dialogue induites par l'incompréhension ou le malentendu. Il s'agit d'une démarche rétrospective et corrective, et non anticipatoire. Avant d'aborder notre méthode, nous présentons le principe général du modèle de dialogue (de type « jeu de dialogue stratégique ») que nous avons utilisé (Caelen, 2003) et quelques définitions :

- Acte de dialogue (Fp) : est un acte de langage comprenant une force illocutoire et un contenu propositionnel, selon la logique illocutoire de (Vanderveken, 1990). Les forces illocutoires retenues dans le domaine du dialogue oral homme-machine sont présentées dans le tableau 2 (Caelen, 2003).

Acte	Signification
F <sup>A</sup>	faire ou exécuter une action (en verbal ou non-verbal)
F <sup>F</sup>	(faire-faire) demander de faire une action à l'allocataire
F <sup>S</sup>	(faire-savoir) communiquer une information
F <sup>FS</sup>	(faire faire-savoir) demander une information
F <sup>P</sup>	(faire pouvoir) donner un choix, faire une invite
F <sup>D</sup>	(faire devoir) obliger sans donner d'alternative

Tableau 2 : Actes de dialogue dans le système Mélima

- But de dialogue : un but est généralement un état du monde ou un état mental que l'on veut atteindre (par exemple obtenir un renseignement, donner une information, etc.). Un but de dialogue est maintenu durant un échange entre un utilisateur et le système (Caelen, Nguyen, 2004), jusqu'à ce qu'il soit satisfait ou abandonné.
- Stratégie de dialogue : la stratégie de dialogue  $\delta$  est la manière de gérer les tours de parole entre l'utilisateur et le système pour conduire efficacement le but de dialogue de l'utilisateur. La stratégie vise à choisir la meilleure direction d'ajustement du but à un moment donné. Nous distinguons les stratégies de dialogue suivantes : directive, réactive, coopérative, de détour (ou constructive) et négociée (Caelen, 2003).

### 3.1. Traitement des malentendus

#### 3.1.1. Malentendus

Les malentendus sont des phénomènes provenant du système de dialogue, mais aussi de l'utilisateur. Ils sont plus faciles à traiter que les incompréhensions, c'est pourquoi nous commençons par eux. Nous considérons les deux cas ci-après :

**M1** : Le malentendu est détecté par le CD comme une contradiction dans sa base de faits  $B_f$  car l'utilisateur a énoncé  $F_{Up}$  dont les effets contredisent un certain  $q \in B_f$ , c'est-à-dire  $p \rightarrow \neg q$ . Dans ce cas, le CD va signaler cette contradiction à l'utilisateur et lui demander un choix ( $p$  ou  $\neg p$ ).

1. Si U maintient  $p$  (U est sûr de lui) alors il s'agit ensuite pour le système de propager ce choix dans sa base de connaissance d'une part puis dans  $B_f$  et dans tous les conséquents de  $p$  (notés  $Q(p)$ ) et de faire les mises à jour nécessaires ; si une nouvelle contradiction apparaît lors de la mise à jour, le processus recommence.
2. Si U s'aperçoit de son erreur alors le dialogue se poursuit sans modification.

**M2** : Le malentendu est détecté par l'utilisateur qui conteste un acte antérieur de la machine  $F_{Mp}$  par  $F_{U\neg q}$ . Cette contestation est soit directe, c'est-à-dire  $q \rightarrow \neg p$ , soit indirecte, c'est-à-dire  $q$  contredit un des conséquents de  $p$ , soit  $\exists p^* \in Q(p) : q \rightarrow \neg p^*$

1. La machine peut alors demander à U de lui expliquer  $q$  tel que  $\neg p \rightarrow q$  qui lui permet de lever son erreur sur  $p$  ou bien,
2. Elle peut tenter de corriger l'erreur elle-même en recherchant les effets d'une de ses actions passées  $F_{Mp}$  indésirable, à travers  $p$  et toutes les conséquents de  $p$  ( $Q(p)$ ) tels que  $p \vee p^* \rightarrow \neg q \mid p^* \in Q(p)$  pour corriger les connaissances ou les faits à propos de  $p$ .

**Exemple 1. Malentendu causé par des erreurs de l'utilisateur**

U : Je voudrais réserver la salle Lafayette  
M : Elle est disponible demain  
U : Entendu, je la prends et prévenez aussi les membres du projet PVE.  
 $F_M^A(\text{réserver}(\text{salle\_Lafayette}) \wedge F_M^A(\text{prévenir}(\text{PVE})), p=\text{prévenir}(\text{PVE}) \rightarrow \text{envoyer}(10 \text{ messages}) \rightarrow \text{organiser\_réunion}(10 \text{ personnes}) \rightarrow q=\text{taille}(\text{salle\_réunion}) \geq 10$   
*Or taille(Lafayette) = 8*  
M : je crois qu'il y a un malentendu car les membres de PVE sont plus de 10 et la salle ne peut accueillir que 8 personnes au maximum

**Exemple 2. Malentendu causé par des erreurs de la machine levées par U**

M : Je réserve la salle pour demain ( $F(\text{salle}, \text{réserver}, 20 \text{ juillet})$ )  
U : Ah non demain je suis en vacances pour une semaine, ce n'est pas possible  
 $F^S(\text{vacances}) \wedge F^S(\text{durée}(\text{vacances})=7) \rightarrow \text{Absent}(U, 20 \text{ juillet}-27 \text{ juillet}) \rightarrow$   
 $\text{Indisponible}(\text{salle}, 20 \text{ juillet}-27 \text{ juillet}) \rightarrow \text{Annuler } F_M^A(\text{salle}, 20 \text{ juillet}) \wedge F_M^{\text{FS}}(\text{date} > 27 \text{ juillet})$   
M : Alors je vous propose de réserver la salle après le 27 juillet. Avez-vous une préférence de date ?

**Exemple 3. Malentendu causé par des erreurs de la machine levées par elle-même**

M : Je réserve la salle pour demain  
U : Est-ce que le technicien sera là ?  
M : La salle sera ouverte, pas de problème  
U : Je ne comprends pas  
U conteste le conséquent  $F_M^S(q)$  pour  $F_U^{\text{FS}}(p)$  avec  $q = \text{ouvrir}(\text{salle})$  et  $p = \text{présent}(\text{technicien})$  donc M cherche dans  $Q(q)$  un autre  $q'$  tel que  $p \rightarrow q' \in Q$  qu'il propose à U par  $F_M^{\text{FS}} q'$   
Dans ce cas la machine doit chercher le malentendu qu'elle a introduit. Il faut comprendre que l'utilisateur associe technicien avec matériel de projection. Donc la machine doit rechercher dans l'ontologie de l'application toutes les attributions d'un technicien dans la tâche. Elle trouve dans cette base que technicien est aussi en charge du matériel, elle rectifie par :  
M : Ne vous inquiétez pas, il s'occupera aussi du matériel de projection  
U : Ah oui d'accord, merci.

**3.1.2. Tactiques de traitement des malentendus**

La stratégie de traitement des malentendus se fonde sur chaque type de malentendu.

1. Si l'utilisateur et le système arrivent à la clôture de dialogue et qu'un malentendu est détecté par le CD, celui-ci propose de terminer le dialogue malgré tout en appliquant la stratégie réactive (en précisant qu'il est trop tard pour traiter ce malentendu) en cas de refus de U, on traitera ce malentendu comme dans le cas suivant (n° 2)

$$\text{but}(\text{Clôture}) \wedge F_U^S q \rightarrow (\delta = \text{réactive}) \wedge F_M^{\text{FS}}(\text{Clôture})$$

2. Dans un échange ordinaire, si le CD détecte un problème, il le signale à l'utilisateur et demande à l'utilisateur s'il maintient son dire en appliquant la stratégie directive :

$$F_U(q) \rightarrow (\delta = \text{directive}) \wedge F_M^S(p) \wedge F_M^S(q \rightarrow \neg p) \wedge F_M^{\text{FS}} q$$

3. Si l'utilisateur détecte un malentendu, alors le système vérifie la cohérence de sa base de faits, la corrige éventuellement ou demande à l'utilisateur des informations complémentaires (ce qui peut guider à trouver cette contradiction) en appliquant la stratégie coopérative. Une suggestion  $q'$  est donc introduite pour guider au préalable l'utilisateur :

$$F_U^S(q) \rightarrow (\delta = \text{coopérative}) \wedge (F_M^S(p) \vee F_M^{\text{FS}}(q'))$$

4. Si l'utilisateur conteste le conséquent  $F_M^S(q)$  pour  $F_U^{\text{FS}}(p)$ , alors la machine va rechercher un autre conséquent  $q'$  tel que  $p \rightarrow q' \mid q' \in Q(p)$ , à l'aide de la stratégie directive pour rectifier le problème :

$$F_U^S(\neg q) \rightarrow (\delta = \text{directive}) \wedge F_M^S(q')$$

5. Si le nombre de malentendus consécutifs  $N_{Mal}$  dépasse un certain seuil  $N_{max}$  (actuellement la valeur du seuil est  $N_{max} = 3$ ), alors le système applique la stratégie directive en utilisant la notification de malentendu :

$$(N_{Incomp} > N_{max}) \wedge F_M(p) \wedge F_U^S(q) \rightarrow (\delta = \text{directive}) \wedge F_M^S(q)$$

### 3.2. Traitement des incompréhensions

#### 3.2.1. Détection par marqueurs dialogiques d'incompréhension (MDI)

Pour marquer les incompréhensions, nous avons introduit une notion de « *Marqueur Dialogique d'Incompréhension – MDI* ». En effet, par définition (voir plus haut) une incompréhension de la machine se manifeste par une structure sémantique erronée ou incomplète, voire vide. L'apparition d'une MDI peut donc être détectée par le CD ce qui lui permet évidemment de la traiter pour trouver des solutions appropriées pour la résoudre. Nous distinguons deux types de MDI :

- MDI-1 : Aucun acte de dialogue  $F_p$  n'est fourni quand l'utilisateur prononce son énoncé : ce cas se produit quand le module de compréhension ne rend aucun schéma sémantique à la sortie (facteur 2.1) ou bien quand tous ses schémas sémantiques sont hors du contexte actuel de l'interpréteur et donc que ce dernier produit un acte vide de l'utilisateur (facteur 2.3). MDI-1 est défini comme un marqueur d'incompréhension complète.
- MDI-2 : Dénote les cas d'incompréhension partielle (facteur 2.2). Dans ce cas, des segments de l'énoncé de l'utilisateur, qui apportent éventuellement des informations, ne sont pas consommés par le module de compréhension. L'interpréteur doit par conséquent déterminer l'apparition de cette situation en analysant la cohérence des actes du système au tour précédent et les actes actuels de l'utilisateur : s'il y a cohérence, il ignore ces morceaux, sinon, il produit le marqueur MDI-2 pour le CD. Ce marqueur est ainsi considéré comme un marqueur d'incompréhension partielle.

En s'appuyant sur ces marqueurs, le CD doit par la suite trouver des solutions adéquates en visant comme objectif de rendre efficace la poursuite du dialogue.

#### 3.2.2. Tactiques de traitement des MDI

Dans le cas où le CD reçoit des marqueurs d'incompréhension MDIs, il peut traiter ces MDIs en appliquant les tactiques illustrées dans le tableau 2.

Tactiques	Exemples	MDI
Confirmation explicite	Voulez-vous la salle Lafayette ? $\delta = \text{réactive}, F^{FS}(\text{Salle}(\text{lafayette}))$	MDI-2
Confirmation implicite	En ce qui concerne l'Aquarium, quelle date voulez-vous réserver ? $\delta = \text{coopérative}, F^S(\text{Salle}(\text{Aquarium})) \wedge F^{FS}(\text{Date}(x))$	MDI-1 MDI-2
Désambiguïsation	Voulez vous le vidéo projecteur Sony ou Philips $\delta = \text{coopérative}, F^P(\text{Matériel}(\text{Sony}, \text{Philips}))$	MDI-1 MDI-2
Suggestion de solution	Je suppose que vous voulez réserver une salle, voulez-vous la salle Aquarium ? $\delta = \text{réactive}, F^{FS}(\text{Salle}(\text{Lafayette}, \text{Aquarium}))$	MDI-1
Demande de répéter	Pourriez-vous répéter quelle salle vous voulez ? $\delta = \text{directive}, F^{FS}(\text{Salle}(x))$	MDI-1
Notification d'incompréhension	Je suis désolée mais je ne comprends pas bien ce que vous avez dit. $\delta = \text{directive}, F^S(\text{Erreur}(\text{imcompréhension}))$	MDI-1

Tableau 2. Tactiques du système devant les MDI

L'initiative de demande de confirmation de l'utilisateur est invoquée dans le cas où il y a un doute concernant son souhait. Envisageons les exemples 2 et 3 :

	Énoncé	Actes de dialogue
S	Quelles sont vos préférences de salle ?	$F^{FS}(\text{Salle}(x))$
U	Je veux la salle <i>Lafayette</i>	$F^S(\text{Salle}(x)) \wedge F^S(\text{MDI-2})$
S	Voulez-vous la salle Lafayette ?	$F^{FS}(\text{Salle}(\textit{lafayette}))$

*Exemple 2. Exemple de confirmation explicite*

L'erreur du module de reconnaissance provoque l'incompréhension partielle pour l'énoncé de l'utilisateur dans l'exemple 2. Il demande donc sa confirmation explicite avec la stratégie réactive afin de continuer à dialoguer.

	Énoncé	Actes de dialogue
S	J'ai la salle aquarium...	$F^S(\text{Salle}(\textit{aquarium}))$
U	Je ne sais pas mais j'ai besoin d' <i>un espace</i> de vingt places	$F^S(\text{Taille}(20)) \wedge F^S(\text{MDI-2})$
S	En ce qui concerne l'Aquarium, quelle date voulez-vous réserver ?	$F^S(\text{Salle}(\textit{aquarium})) \wedge F^{FS}(\text{Date}(x))$

*Exemple 3. Exemple de confirmation implicite*

Dans l'exemple 3, devant l'incompréhension partielle causée au niveau lexical, le système doit nécessairement appliquer la tactique de demander la confirmation implicite de l'utilisateur en appliquant la stratégie coopérative : s'il ne la réfute pas, alors le système considère cette information comme sûre et continue le traitement.

	Énoncé	Actes de dialogue
S	Quelle salle voulez-vous ?	$F^{FS}(\text{Salle}(x))$
U	La dernière <b>sale</b> <la dernière salle> (ou la salle pour dix personnes)	$F^S(\text{MDI-1})$ (ou $F^S(\text{Taille}(>10)) \wedge F^S(\text{MDI-2})$ )
S	Voulez-vous la salle Lafayette ou Aquarium ?	$F^P(\text{Salle}(\textit{aquarium}, \textit{Lafayette}))$

*Exemple 4. Tactique de désambiguïsation*

La désambiguïsation, comme illustrée par l'exemple 4, peut être utilisée avec les marqueurs MDI-1 et MDI-2 tandis que la suggestion, présentée dans l'exemple 5, n'est utilisée que dans le cas des incompréhensions complètes, de manière à clarifier le dialogue et la tâche. Supposons, par exemple, que les deux salles concernées ont plus de dix places dans l'exemple 4.

	Énoncé	Actes de dialogue
S	Bonjour monsieur Dupont, je vous écoute ...	$F^S(\text{Nom}(\textit{Dupont})) \dots$
U	Je veux <b>une sale</b> <je veux une salle>	$F^S(\text{MDI-1})$
S	Je suppose que vous voulez réserver une salle, voulez-vous la salle Aquarium ?	$F^{FS}(\text{réservé}(x))$ and $F^S(\text{Salle}(\textit{Aquarium}))$

*Exemple 5. Tactique de suggestion*

Les deux dernières tactiques – demande de répéter et notification d'incompréhension – ne sont appliquées que dans le cas où aucune information n'est trouvée par les deux modules de compréhension et d'interprétation. Dans ce cas, la stratégie directive doit être utilisée, afin de conduire l'utilisateur à continuer le dialogue. La notification d'incompréhension peut être présentée comme l'exemple 6.

	Énoncé	Actes de dialogue
S	Quelles sont vos préférences de salle ?	$F^{FS}(\text{Salle}(x))...$
U	J'en sais rien moi...	$F^S(\text{MDI-1})$
S	Désolé mais je ne comprends pas ce que vous dites.	$F^S(\text{Incompréhension})$

*Exemple 6. Notification d'incompréhension*

### 3.2.3. Tactiques de traitement des incompréhensions

La stratégie de traitement des erreurs est fondée sur les marqueurs dialogiques d'incompréhension, ainsi que sur les calculs des attentes de l'utilisateur. Nous distinguons les attitudes suivantes portant sur l'état du système par des MDI :

- Si l'utilisateur et le système arrivent au thème de clôture de dialogue et qu'un MDI est signalé au CD (soit MDI-1, soit MDI-2), le CD manifeste le souhait de terminer le dialogue en appliquant la stratégie réactive :

$$\text{but}(\text{Clôture}) \wedge \text{MDI} \rightarrow (\delta = \text{réactive}) \wedge F^S_M(\text{Clôture})$$

- Si le but de dialogue actuel  $b$  est atteint ou satisfait, le système pose un nouveau but en demandant une confirmation implicite et en appliquant la stratégie coopérative pour une incompréhension MDI-2 et directive pour MDI-1 :

$$(b = \text{atteint}) \wedge \text{MDI-2} \rightarrow (\delta = \text{coopérative}) \wedge F^S_M(b) \wedge F^{FS}_M(?b)$$

$$(b = \text{atteint}) \wedge \text{MDI-1} \rightarrow (\delta = \text{directive}) \wedge F^S_{M\pi}(b) \wedge F^{FS}_M(?b)$$

- Si le système est en train de demander une information (le but actuel  $b$  n'est pas encore atteint), alors le système applique la stratégie réactive en essayant de rattacher à ce but des données qui ne sont pas encore consommées par les deux modules de compréhension et d'interprétation. Une suggestion est donc introduite pour guider l'utilisateur :

$$F^{FS}_M(p) \wedge \text{MDI} \rightarrow (\delta = \text{réactive}) \wedge F^{FS}_M(q)$$

donc  $q = \text{inférer}(b, \text{MDI})$  (le calcul de  $q$  se fondant sur le but actuel et MDI)

- Si le système est en train de donner une information et que le but actuel n'est pas encore atteint, alors il informe l'état actuel de ce but en cas d'incompréhension :

$$F^S_M(p) \wedge \text{MDI} \rightarrow (\delta = \text{coopérative}) \wedge F^S_M(b)$$

- Si le nombre d'incompréhensions consécutives  $N_{\text{Incomp}}$  dépasse un certain seuil  $N_{\text{max}}$  (actuellement la valeur du seuil est  $N_{\text{max}} = 3$ ), alors le système applique la stratégie directive en utilisant une demande de répétition ou la notification d'incompréhension :

$$(N_{\text{Incomp}} > N_{\text{max}}) \wedge F_M(p) \wedge \text{MDI} \rightarrow (\delta = \text{directive}) \wedge F^S_M(\text{MDI}) \wedge F_M(p)$$

## 4. Expérimentation

L'implémentation liée à ces divers traitements a été mise en œuvre dans le système de dialogue Méлина développé au laboratoire CLIPS (Nguyen, 2005). Mais aucune évaluation systématique n'a été faite encore à ce jour avec des sujets. Un exemple de dialogue est explicité ci-après. Au cours de ce dialogue, des erreurs du module de reconnaissance de la parole provoquent des incompréhensions et un malentendu entre le nombre des participants et la taille de la salle. L'énoncé entre parenthèses est celui qui est prononcé par l'utilisateur et l'énoncé en gras est le résultat fourni par le module de reconnaissance de la parole.

Énoncés et actes de dialogue	
S0	<i>FS</i> [Ouverture (commencer)] ici le service de réservation du laboratoire clips, je suis Mélina et vous ?
U1	<b>je si elle vient</b> <je suis nguyen> <i>FS</i> [Incomprehension (MDI-1)]
S1	<i>FS</i> [Ouverture (répéter)] & <i>FFS</i> [Usager(x)] je suis Mélina ! puis-je vous demander votre nom ?
U2	nguyen <i>FS</i> [Nom (nguyen)]
S2	<i>FS</i> [Groupe (geod)] & <i>FFS</i> [Reservation(x)] quelles sont vos préférences de réservation ?
U3	j'aimerais réserver la salle lafayette <i>FF</i> [Action (réservé)] & <i>FS</i> [NomSalle (lafayette)]
S3	<i>FS</i> [Taille(45)] & <i>FFS</i> [Temps(x)] c'est une salle de 20 places, c'est pour quel jour, s'il vous plaît ?
U4	<b>est-elle disponible le genre jeudi prochain</b> <est-ce qu'elle est disponible jeudi prochain> <i>FS</i> [Date(20051110)]
S4	<i>FFS</i> [Heure(x)] à quelle heure s'il vous plaît ?
U5	<b>non de dix heures à midi</b> <ben de dix heures à midi> <i>FS</i> [Heure(600)] & <i>FS</i> [Duration(120)]
S5	<i>FS</i> [Reservation(2)] je vous réserve la salle lafayette à 10 heures pour jeudi 10 novembre 2005 pendant 2 heures
U6	<b>j' aimerais aussi sinon vidéo projecteur</b> <j'aimerais aussi un vidéo projecteur> <i>FS</i> [Materiel (vidéo projecteur)]
S6	<i>FFS</i> [Convocation(x)] voulez-vous que je prévienne les participants ?
U7	<b>oui tous membres projet pve</b> <oui tous les membres du projet pve> <i>FS</i> [Convocation(OK)] & <i>FS</i> [Membre(pve)]
S7	<b>Je suis désolé, mais je crois qu'il y a un malentendu ici. La salle Lafayette n'a que 20 places et les membres du projet PVE sont 40.</b> <i>FS</i> [Malentendu(TailleSalle)] & <i>FS</i> [Taille(20)] & <i>FS</i> [Membres(40)]
U8	<b>Ah bon préviens seulement dix permanents pour cette réunion</b> <Ah bon préviens seulement dix permanents pour cette réunion> <i>FS</i> [Membres(10 permanents)]
S8	<i>FS</i> [Membres(OK)] & <i>FFS</i> [Mode(x)] Entendu, et par email ou par téléphone ?
U9	<b>pas téléphone</b> <par téléphone> <i>FS</i> [Incomprehension (MDI-1)]
S9	<i>FFS</i> [Mode(téléphone)] je vais les contacter par quel moyen, par téléphone ?
U10	oui par téléphone <i>FS</i> [Mode(téléphone)]
S10	<i>FS</i> [Convocation(OK)] très bien, je vais prévenir les participants
U11	merci bien au revoir <i>FS</i> [Cloture(x)]
S11	<i>FS</i> [Cloture(OK)] au revoir monsieur nguyen à bientôt.

*Exemple 7. Un dialogue oral enregistré par Mélina avec des incompréhensions*

L'incompréhension est effectivement détectée à la fois par le module de compréhension et d'interprétation en distinguant deux marqueurs dialogiques d'incompréhension : MDI-1 signifie l'incompréhension complète et MDI-2 la partielle. Les tactiques du système sont ici : demande de confirmation explicite et implicite, désambiguïsation, suggestion, demande de répéter, notification d'incompréhension. Par conséquent, une fois que l'incompréhension est signalée par ces marqueurs, le contrôleur du dialogue peut essayer de trouver la solution adéquate, afin d'assurer l'avancement efficace du dialogue.

En ce qui concerne le malentendu, Mélina est capable de traiter le problème dans ce dialogue, représenté dans les tours de parole (U7, S7). Dans ce cas, ce malentendu a été réglé par l'utilisateur en réduisant le nombre de participants pour s'adapter à la taille de la salle demandée.

## 5. Conclusion

Les erreurs au cours d'un dialogue sont des phénomènes que subit la plupart des systèmes de dialogue. Nous avons abordé, dans cet article, une approche pour traiter non seulement des incompréhensions, mais également des malentendus au niveau du contrôleur de dialogue (CD). Deux marqueurs MDI-1 et MDI-2 ont été proposés afin de détecter et de signaler les types d'incompréhensions au CD afin de les traiter de manière plus spécifique. En ce qui concerne les malentendus, la stratégie est d'attendre une contradiction patente puis de la traiter en remontant à la source du problème et en répercutant les mises à jours dans la base de faits et/ou de connaissance du système. En appliquant cette approche au système de dialogue Mélina, nous avons obtenu quelques premiers résultats prometteurs.

## Références

- BERNSEN N.O., DYBKJAER H., DYBKJAER L. (1998). *Designing Interactive Speech Systems: from first ideas to user testing*. Springer Verlag.
- CARPENTER P., JIN C., WILSON D., ZHANG R., BOHUS D., RUDNICKY A. (2001) « Is This Conversation on Track? », in *Proceedings of Eurospeech 2001*. Aalborg.
- CAELEN J. (2003). « Stratégies de dialogue ». In *Actes de MFI'03 (Modèles Formels de l'Interaction)*. CEPADUES, Lille.
- CAELEN J., NGUYEN H. (2004). « Gestion de buts de dialogue ». In *Actes de TALN 2004*. Fès : 345-350.
- HAZEN T.-J., BURIANEK T., POLIFRONI J., SENEFF S. (2002). « Recognition Confidence Scoring for Use in Speech Understanding Systems ». In *Computer Speech and Language* 16 (1) : 49-67.
- NGUYEN H. (2005). *Dialogue Homme-Machine : Modélisation de multisession*. Thèse de l'Université Joseph Fourier, Grenoble.
- RUDNICKY A., THAYER E., CONSTANTINIDES P., TCHOU C., STERN R., LENZO K., XU W., OH A. (1999). « Creating natural dialogs in the Carnegie Mellon Communicator System ». In *Proceedings of Eurospeech* : 1531-1534.
- VANDERVEKEN D. (1990). *La logique illocutoire*. Mardaga, Bruxelles.
- WALKER M., WRIGHT J., LANGKILDE I. (2000). « Using Natural Language Processing and Discourse Features to Identify Understanding Errors in a Spoken Dialogue System ». In *Proceedings of the 17th International Conference of Machine Learning* : 1111-1118.
- YANKOLOVICH N., LEVOW G.A., MARX M. (1995). « Designing SpeechActs: Issues in Speech User Interfaces ». In *Proceedings of CHI'95* : 369-376.