

# Deux indicateurs de la collaboration de l'utilisateur en dialogue homme-machine : une étude expérimentale

Virginie Amiel<sup>1,2</sup> - Ludovic Le Bigot<sup>1</sup> – Patrice Terrier<sup>2</sup> –

Jean-Marie Cellier<sup>2</sup> – Lisa-Marie Babin<sup>1,2</sup>

<sup>1</sup> France Télécom – Recherche et Développement  
2, avenue P. Marzin – 22300 Lannion  
[virginie.amiel@francetelecom.com](mailto:virginie.amiel@francetelecom.com)

<sup>2</sup> UMR 5551 - Université Toulouse II  
5, allée A. Machado – 31058 Toulouse Cedex 9  
[terrier@univ-tlse2.fr](mailto:terrier@univ-tlse2.fr)

## RESUME

Comment peut se manifester la collaboration de l'utilisateur en situation d'interaction avec un système de dialogue en langue naturelle ? A partir des travaux sur la construction du contexte partagé dans les interactions langagières, deux comportements collaboratifs caractéristiques ont été isolés : la correspondance des réponses aux attentes exprimées par le partenaire et la reprise de la longueur et de la structure de ces énoncés. Ils ont été étudiés dans une expérimentation avec un service réel, lors de tâches de recherche d'informations à base de scénarii, avec deux modes de dialogue (oral vs écrit). Les résultats ont montré que les utilisateurs avaient des comportements plus collaboratifs à l'oral qu'à l'écrit lors des interactions avec ce service. Pour finir, des perspectives pour l'analyse et l'amélioration de la collaboration dans le dialogue homme-machine ont été proposées.

**MOTS CLES** : Dialogue, IHM, Collaboration, Construction du contexte partagé.

## ABSTRACT

What is the expression of user collaboration when interacting with a natural language dialogue system? Based on studies dealing with the construction of shared context in linguistic interactions, two specific co-operative behaviours were isolated: correspondence between answers and partner expectations, reuse of partner utterances structure and length. They were studied hereby on a real system experimentation, using information retrieval tasks based on scenarios with two dialogue modes (written vs vocal). Results showed that user behaviours were more co-operative using the vocal mode while interacting with this service. Eventually, analysis and improvement perspectives of collaboration within man-machine dialogue were proposed.

**KEYWORDS** : Dialogue, HCI, Collaboration, Grounding.

## INTRODUCTION

Les avancées technologiques de ces dernières années ont permis l'émergence de systèmes de dialogue naturel<sup>1</sup> permettant des interactions verbales dans le cadre de tâches de recherche et de consultation d'informations, d'assistance dans des situations personnelles ou professionnelles ou d'aide à la prise de décision. La conception de ces systèmes dialoguants est basée sur des modèles linguistiques et informatiques de plus en plus sophistiqués dans lesquels la prise en compte de l'utilisateur reste insuffisante. L'utilisateur est encore envisagé comme celui qui va se servir du système, et non comme faisant partie intégrante de l'interaction. Pourtant, dans les situations de communication orientées par un but, ce n'est pas seulement chaque acte individuel qui est orienté par ce but mais l'ensemble de l'interaction [18], ce qui signifie que le dialogue émerge de l'interaction entre l'utilisateur et le système dialoguant. En conséquence, pour prendre en compte l'utilisateur, les concepteurs ont besoin non seulement de connaissances sur ses capacités cognitives, mais aussi (et peut-être surtout) de connaissances sur les processus spécifiques au dialogue homme-machine (DHM). Ainsi, des auteurs [4, 20] ont attribué au processus de construction du terrain commun un rôle déterminant dans le dialogue en le considérant comme le principal *processus actif de la collaboration* [1, 10]. Alors que de nombreuses études sont réalisées pour modéliser la collaboration en DHM (e.g., [27]), trop peu d'études tentent d'isoler et d'analyser les manifestations de la collaboration des utilisateurs. Dans ce cadre, ce papier présente une expérimentation, avec un service réel de recherche d'information en dialogue naturel à l'oral et à l'écrit, qui se centre sur la collaboration entre l'utilisateur et le système.

## LA COLLABORATION EN DHM

De nombreuses recherches en psychologie portant sur les dialogues homme-machine se sont inspirées des modèles de la communication humaine [10]. Par

---

<sup>1</sup> Pour une revue sur les systèmes vocaux, voir [23]

exemple, Brennan (e.g. [4, 6]) considère qu'il est possible d'étudier les échanges conversationnels homme-machine comme des dialogues interprétables et comparables suivant les théories habituelles de la communication. Concernant les dialogues homme-homme (DHH), Clark et Wilkes-Gibbs [15] ont étudié la collaboration dans le cadre de l'établissement de références communes. Ils concluent que la collaboration est nécessaire pour que les partenaires établissent mutuellement le sens de ce que chaque locuteur dit.

La collaboration a été définie dans le dictionnaire (Le Robert) à la fois comme l'action de travailler en commun (avec quelqu'un) et comme le résultat de cette action. Le terme coopération est donné comme synonyme. Au contraire, Allwood *et al.* [1] placent la collaboration comme une dimension constitutive de la coopération (la seconde dimension étant la coordination). En psychologie, la différence entre coopération et collaboration peut être située au niveau de la division du travail : dans la coopération, les partenaires résolvent des sous-tâches individuelles et ensuite rassemblent les résultats partiels dans une sortie finale, alors que dans la collaboration, les partenaires font le travail ensemble [16]. Pour Clark et Schaefer [14], un acte collectif est un acte réalisé par un collectif, où deux personnes au minimum agissent ensemble. Il peut être décomposé en trois actes distincts : un acte collectif de A et B et deux actes individuels, un de A et un de B ([14], p.263). L'acte collectif implique la collaboration entre A et B, alors que les actes individuels impliquent la coordination<sup>2</sup>. Pour Clark et ses collaborateurs, la collaboration est centrale pour construire le dialogue [15]. Elle implique non seulement que le locuteur produise des phrases compréhensibles et que l'auditeur comprenne ces phrases, mais aussi que les deux interlocuteurs établissent la croyance mutuelle que le destinataire a compris ce que le locuteur voulait signifier. Ce processus est appelé *grounding*, il consiste pour chaque partenaire à la fois à spécifier un contenu et à l'incorporer dans le terrain commun (ce point est développé dans la section suivante).

En ce qui concerne le DHM, des auteurs [27] ont constaté que les utilisateurs étaient plus satisfaits au niveau subjectif lorsqu'ils percevaient le système comme collaboratif. Au niveau des interactions, les avis sur la collaboration entre les utilisateurs et les systèmes divergent. Amalberti *et al.* [2] avancent que les utilisateurs ne sont pas collaboratifs en DHM, car ils ont tendance à utiliser le système comme un outil plutôt que comme un partenaire pour la résolution *collective* d'une tâche, alors que d'autres auteurs [20] affirment que les

<sup>2</sup> Cette coordination a été étudiée en dégageant deux niveaux : N1 où se fait la synchronisation des vocalisations avec l'attention de l'auditeur et N2 où se fait la synchronisation de la présentation des expressions du locuteur avec les expressions de son auditeur [11].

utilisateurs sont collaboratifs au moins au début du dialogue. Les systèmes émergents de dialogue naturel sont conçus avec des capacités identifiées pour augmenter la convivialité des interactions (en s'appuyant sur des capacités comme la négociation, la flexibilité de l'interaction, l'adéquation des formes de réponses, pour plus de détails consulter [24, 25]). En considérant, à la suite de Dillenbourg [16], qu'un paramètre important de la collaboration est la situation de communication, les situations de dialogue avec de tels services automatiques peuvent être qualifiées de collaboratives car elles se caractérisent par une symétrie des possibilités des partenaires dans l'interaction, des buts communs et « l'interactivité, la synchronicité et la négociabilité » ([16], p. 8). Mais face à de tels systèmes, comment se comporte effectivement l'utilisateur ? Quels sont les comportements de collaboration utilisés en DHM ?

### LA CONSTRUCTION DU CONTEXTE PARTAGE

Des travaux [4, 20] ont établi que le processus de construction du contexte partagé<sup>3</sup>, i.e. *le grounding*, était déterminant en DHM, dans le sens où il serait le principal « processus actif » de la collaboration [1, 10, 15]. Le *grounding* a été défini comme le processus permettant l'élaboration d'une représentation de « *la somme des connaissances mutuelles des interlocuteurs, de leurs croyances et de leurs suppositions* » ([10], p.93). A la suite de l'étude d'Amalberti *et al.* [2], Johnstone *et al.* [20] ont indiqué que les différences entre DHH et DHM résidaient principalement dans l'utilisation de certains mots et d'échanges typiques utilisés dans le *grounding* et les clôtures (e.g., « au revoir »).

Le *grounding* a été décrit en considérant la communication comme une forme d'action collaborative où les partenaires cherchent des indices de la compréhension chez les autres et fournissent des indices de leur propre compréhension [7]. Pour Clark et Schaefer [13, 14], la construction du terrain commun se déroule dans les contributions sur la base de l'échange de preuves conditionnelles de la compréhension des partenaires. Ces preuves conditionnelles peuvent être positives, l'élément est alors compris, accepté et partagé, ou négatives, l'élément nécessite alors un sous-dialogue de clarification pour être partagé. Les preuves positives, qui constituent les types de retours habituellement recherchés par le partenaire, peuvent prendre plusieurs formes dans la conversation humaine [12, 14]. Dans les DHM, les indicateurs de compréhension traduisant le *grounding* par l'utilisateur sont moins nombreux [7]. Ainsi, de nombreuses difficultés pourraient être en partie expliquées par « *le problème de la construction du terrain commun dans les IHM* » (notre traduction [7], p.3). Cependant, malgré ce manque d'indices, le

<sup>3</sup> Le terme de « shared context » est utilisé pour les DHM [6], alors que pour les DHH on utilise « common ground » [11].

grounding est à l'œuvre du côté de l'utilisateur : qu'ils communiquent avec ou à travers un ordinateur, que l'interaction se fasse à travers des icônes, du texte ou de la parole, les gens s'adressent à un partenaire *interactif* [4].

Le grounding se manifeste dans les DHM par différents types de comportements verbaux. Les preuves de la compréhension, dans la conversation humaine, peuvent prendre trois formes principales : (1) une attention continue de la part du récepteur, (2) des signes de reconnaissance, ou (3) un tour de parole pertinent par rapport à ce qui vient d'être dit [12]. En DHM, du fait que seul le matériel verbal des interactions est considéré, c'est principalement dans le tour de parole suivant que les preuves de la compréhension sont à chercher. Ainsi, l'utilisateur peut donner au système soit une réponse qui correspond plus ou moins aux attentes exprimées dans l'énoncé système précédent, soit une information centrée sur sa tâche. Par ces comportements l'utilisateur pourrait être qualifié comme plus ou moins collaboratif. De plus, des études ([21] en CMC<sup>4</sup> ; [5]) ont permis de mettre en évidence des indices conversationnels de grounding. Par exemple, McCarthy et Monk [21] ont isolé des variables pertinentes pour l'étude des conversations. Au niveau de l'analyse quantitative du contenu des conversations, les auteurs proposent d'étudier les pronoms personnels (à la suite de [4]) et les expressions co-référentielles « it, that, they, she » se référant à un groupe nominal dans le tour précédent. Concernant les mesures du terrain commun et de l'orientation mutuelle<sup>5</sup>, ils proposent d'étudier les références au sujet d'un message au prochain tour disponible et l'entraînement syntaxique, qui réfère au fait que la syntaxe des énoncés reflète celle des énoncés du partenaire en terme de reprise de la longueur et de la structure. De même, Brennan [5] a montré que les individus ont tendance à utiliser le même langage que leur interlocuteur : qu'ils conversent avec une personne ou avec un ordinateur, à l'écrit ou à l'oral, ils tendent à converser avec leur partenaire dans les termes qu'il utilise. Plusieurs hypothèses ont été proposées pour expliquer ce phénomène. Pour les auteurs précédemment cités, il constitue une manifestation du grounding, proche de l'entraînement syntaxique dans les DHH [8, 22], alors que d'un point de vue psycholinguistique, les convergences des faits de langue sont interprétés en termes d'amorçages automatiques [19]. Ainsi, la reprise du lexique du partenaire au cours de l'interaction serait un indicateur important de l'adaptation au partenaire et donc de la collaboration en DHM.

## OBJECTIF DE L'ETUDE

La collaboration des utilisateurs dans les DHM se manifeste donc principalement par la construction du

contexte partagé. Elle peut être visible dans les contributions au moins au niveau des réponses aux attentes du partenaire et au niveau de la convergence des énoncés des différents interlocuteurs. De nombreux travaux sur l'utilisation du langage ont souligné aussi la nécessité d'étudier les DHM à l'oral et à l'écrit (e.g., [17]). De plus, Clark et Brennan [12] ont montré que le mode d'interaction avait une influence sur les techniques de grounding. L'hypothèse est que la collaboration est présente en DHM, et qu'elle se manifeste dans les comportements de l'utilisateur en termes de correspondance des réponses aux attentes du partenaire et en termes de reprise de la longueur et de la structure des énoncés de l'interlocuteur. Cependant, les manifestations de la collaboration peuvent être différentes en fonction de la modalité d'interaction. Plus précisément, une interaction à l'oral devrait favoriser l'apparition d'indicateurs de collaboration par rapport à une interaction à l'écrit.

## EXPERIMENTATION

Une expérimentation a été conduite pour étudier des manifestations de la collaboration d'utilisateurs en situation d'interaction avec système réel de dialogue en langage naturel à l'oral et à l'écrit.

**Participants.** Les participants étaient au nombre de 24 (22 hommes, 2 femmes), l'âge moyen était de 23 ans ( $E.T.=2.01$ ). Le niveau d'étude était élevé (en moyenne 4.4 années d'étude après le bac,  $E.T.=1.28$ ). Les participants plutôt familiarisés avec l'informatique, se sont définis eux-mêmes avec un niveau en informatique moyen ( $M=3.55$   $E.T.=0.93$ , sur une échelle en 5 points de très faible à fort), et comme utilisant régulièrement l'outil informatique (courrier électronique, Internet et logiciels bureautiques). Par contre, aucun n'employait régulièrement des services interactifs en ligne ou par téléphone. Les participants ont été répartis de manière aléatoire dans les deux groupes expérimentaux.

**Matériel et procédure.** Le service utilisé *ArtimisVoyage* est un démonstrateur de consultation d'horaires de trains et/ou d'avion (à base de dialogue intelligent en langage naturel). L'utilisateur effectuait une recherche avec plusieurs critères : villes de départ et d'arrivée, jour et heure de départ et/ou d'arrivée, moyen de transport. Le système lui proposait en retour les solutions correspondant ou s'approchant au mieux de sa requête. Une des caractéristiques principales était qu'*ArtimisVoyage* assistait fortement l'utilisateur pour la recherche. Le service fonctionnait à l'oral par téléphone (avec synthèse et reconnaissance vocale) et à l'écrit avec une interface de type Web. L'unique moyen de communication avec le système était l'utilisation du langage. Le contenu des énoncés systèmes était identique pour les deux modalités. Un énoncé utilisateur à l'oral ou à l'écrit produisait exactement la même réponse du système. Deux phases pouvaient décrire une

<sup>4</sup> Computer Mediated Communication

<sup>5</sup> L'orientation mutuelle est pour « un processus social subtil qui peut laisser présager une difficulté quand le système dégrade la communication » (notre traduction, [21] p° 317).

recherche typique avec *ArtimisVoyage* : (1) une phase d'émission de critères, et (2) une phase d'affinage des solutions (cf. Figure 1 pour un exemple de dialogue).

Dialogue oral	
	Systeme
1	- Bienvenue sur ARTIMIS-Voyage. [...] Que puis-je pour vous ?
	Utilisateur
2	- Je desire un Paris - Toulouse
	Systeme
3	- Un trajet de Paris pour Toulouse. Vous voulez partir quel jour ?
	Utilisateur
4	- ben j'aimerais partir demain matin
	Systeme
5	- Un trajet le matin de Paris pour Toulouse le vendredi 6 juillet 2001. Quel moyen de transport ? En train ou en avion ?
	Utilisateur
6	- en train
	Systeme
7	- Veuillez patienter s'il vous plaît, je consulte les agents. - Il y a un trajet en train partant de Paris le vendredi 6 juillet 2001 à 7 heures 20 et arrivant à Toulouse à 13 heures. Voulez-vous des détails, les alternatives ou une autre information ?
	Utilisateur
8	- détails
	Systeme
9	- Il y a deux étapes. Pour la première étape, [...]. Pour la 2eme étape, [...]. Voulez-vous les alternatives ou une autre information ?
	Utilisateur
10	- les alternatives
	Systeme
11	- Il y a trois alternatives. Le trajet le plus rapide en train [...] Voulez-vous le trajet alternatif suivant ou une autre information ?
	Utilisateur
12	- Non, ça va merci
	[Fin]

Figure 1 : Exemple de dialogue transcrit à l'oral.

*Scénarii de dialogue.* Des tâches de recherche d'informations sous forme de scénarii ont été fournies aux utilisateurs. Cinq critères étaient donnés (villes de départ et d'arrivée, jour et heure de départ, moyen de transport) pour effectuer la recherche (Figure 2). Quatre types de tâches étaient demandées à l'utilisateur, soit : (a) une consultation d'horaire, (b) un nombre d'horaires correspondant à la demande, (c) le nombre d'étapes du trajet, (d) le trajet le plus rapide pour réaliser la tâche (le nombre de tours de parole théorique était identique pour toutes les recherches). L'ordre des critères était fixe. En manipulant les critères et le type de tâche, 8 scénarii simples ont été créés (2 par type de tâche). Ces 8 scénarii ont été répartis par quatre en deux sessions (S1 & S2). L'ordre a été contrebalancé.

Après un recueil expérimental, les échanges entre l'utilisateur et le système ont été transcrits mots à mots. Les énoncés réels des utilisateurs ont été considérés pour l'analyse (et non les énoncés issus des modules de reconnaissance vocale et d'interprétation). Ensuite, les dialogues ont été annotés.

- Départ : Lannion
- Arrivée : Quimper
- Date : demain
- Heure de départ : vers 7 heures
- Transport : train
Combien d'étapes comporte ce trajet ?

Figure 2 : Exemple de scénario pour *ArtimisVoyage*.

**Variables indépendantes.** Les facteurs manipulés étaient l'ordre des modalités (oral puis écrit, écrit puis oral) en intergroupe et la session (S1, S2) en intragroupe. En d'autres termes, douze utilisateurs ont commencé par la modalité orale en session 1 et ont poursuivi par la modalité écrite en session 2, les douze autres ont fait l'inverse.

**Données recueillies.** Les annotations concernaient deux manifestations de la collaboration de l'utilisateur : (1) la nature plus ou moins collaborative de ses réponses et (2) ses reprises de la longueur et de la structure des énoncés de commandes.

(1) La correspondance des réponses aux attentes du partenaire système ou les prises d'initiative étaient surtout présentes dans la première partie des dialogues, i.e. la phase d'émission des critères. Un indice *Énoncés Collaboratifs* a été élaboré pour qualifier les énoncés des utilisateurs par rapport aux attentes exprimées dans les énoncés du système. Cet indice a été calculé en codant l'appariement vs non-appariement de chaque énoncé de l'utilisateur aux demandes du système, pour les tours de parole efficaces<sup>6</sup>. Par exemple, si le système demandait de préciser la ville de départ et que l'utilisateur précisait la ville de départ, alors l'énoncé était codé comme collaboratif. A l'inverse, si l'utilisateur fournissait un jour ou une heure, l'énoncé était codé comme non-collaboratif. Il faut remarquer que les premiers tours de parole ont été codés différemment car ils ne pouvaient pas correspondre à une des catégories. Afin d'obtenir une variable rendant compte de la collaboration de l'utilisateur, la transformation suivante a été réalisée :  $[\ln(\text{enoncecoll}+1) - \ln(\text{enoncecoll}+1)]$ . Elle correspond au logarithme naturel du rapport 'énoncé collaboratif plus un' par 'énoncé non collaboratif plus

<sup>6</sup> Le nombre de tours de parole efficaces correspondait au nombre d'énoncés dédiés à la résolution de la tâche (incluant les énoncés dédiés à l'émission des critères et les énoncés dédiés aux commandes, mais en enlevant les énoncés dus à des erreurs de reconnaissance vocale ou d'orthographe).

un' (il faut noter que les calculs pour cet indice ont été effectués par session sur 4 dialogues). Cette transformation a permis d'exprimer la relation entre les énoncés collaboratifs et non-collaboratifs (échelle de rapport) tout en respectant les différentes conditions d'applications de l'analyse de la variance (ANOVA). En considérant les propriétés des logarithmes, si l'indice obtenu était positif, le sujet pouvait être considéré plutôt collaboratif, et inversement s'il était négatif, le sujet pouvait être considéré plutôt centré sur la tâche.

(2) Les reprises de la longueur et de la structure des énoncés du système par l'utilisateur étaient surtout présentes dans la phase d'affinage des solutions. Un indice *Reprise des énoncés de commande* a été calculé afin de refléter l'adaptation de l'utilisateur au système par le partage des énoncés de commandes. Il a été mesuré pour chaque dialogue en regroupant les commandes selon quatre niveaux : (1) Reprises littérales (présence de tous les mots constituant la commande), (2) Reprises partielles (deux mots), (3) Mots uniques (un mot de la commande) et (4) Reformulations ; puis en les divisant par le nombre de tours de parole efficaces pour la tâche<sup>6</sup>.

## RESULTATS

**Analyse des données.** Les données ont fait l'objet d'une ANOVA pour plan mixte avec le facteur Ordre des modalités en intergroupe et le facteur Session en intragroupe pour les Enoncés collaboratifs. Le facteur Forme des énoncés a été ajouté en intragroupe pour les Reprises des énoncés de commande. Les effets d'interaction ont été étudiés avec des comparaisons planifiées. Les pré-analyses des données sur les scénarii et les tâches n'ont pas montré de différence. De plus, le taux de réussite global était proche de 95%. Les moyennes et les écarts-types ont été reportés dans le tableau (Figure 3).

(1) Pour l'indice Enoncés Collaboratifs, l'analyse n'a pas montré d'effet principal de l'Ordre des modalités,  $F(1,22)=1.50, p>.1$ , ni de la Session,  $F(1,22)=1.03, p>.1$ . Cependant, l'interaction entre les facteurs Session  $\times$  Ordre était significative,  $F(1,22)=25.17, p<.0001$ . Cette interaction reflétait un effet de la modalité. Plus précisément, le nombre d'énoncés collaboratifs était plus important en mode oral qu'en mode écrit en session 1,  $F(1,22)=12.31, p<.01$ , pour le groupe ayant commencé à l'oral puis continué à l'écrit,  $F(1,22)=8.02, p<.01$ , et pour le groupe ayant commencé à l'écrit puis continué à l'oral,  $F(1,22)=18.20, p<.0001$ .

(2) Pour l'indice de Reprise des énoncés de commande, le facteur Forme des énoncés de commandes correspondait : aux Reprises littérales (1) et aux reprises Mots uniques (3). Les Formes (2) et (4) n'ont pas été analysées ici car elles ont été peu fréquentes. L'analyse n'a pas montré d'effet principal de l'Ordre des modalités,  $F(1,22)=.78, p>.1$ , de la Session,

$F(1,22)=1.82, p>.1$ , et de la Forme des reprises,  $F(1,22)=.53, p>.1$ . Les interactions entre les facteurs Ordre  $\times$  Session et Ordre  $\times$  Forme étaient significatives, respectivement  $F(1,22)=16.39, p<.001$ ,  $F(1,22)=4.96, p<.05$ . Cependant, l'interaction entre les trois facteurs était significative,  $F(1,22)=10.08, p<.01$ . Cette dernière reflétait un effet de modalité en session 1. Plus précisément en session 1, le nombre de reprises Littérales de commandes était plus important en mode oral qu'en mode écrit,  $F(1,22)=4.89, p<.05$  ; à l'inverse le nombre de reprises Mots uniques était plus important en mode écrit qu'en mode oral,  $F(1,22)=18.79, p<.0001$ . De la même façon, en mode oral en session 1, le nombre de reprises Littérales de commandes était plus important que le nombre de reprises Mots uniques,  $F(1,22)=5.00, p<.05$  ; à l'inverse, en mode écrit le nombre de reprises Mots uniques était plus important dans le mode écrit que dans le mode oral,  $F(1,22)=8.45, p<.01$ . En session 2, aucune comparaison n'a atteint le seuil de significativité statistique.

Mesures	N	Session 1		Session 2	
		m	E.T.	m	E.T.
<b>(1) Enoncés Collaboratifs</b>					
oral puis écrit	12	1.71	0.91	0.99	0.88
écrit puis oral	12	0.41	0.89	1.50	0.88
Total	24	1.06	1.10	1.24	0.90
<b>(2) Reprises Littérales des énoncés de commande</b>					
oral puis écrit	12	0.79	0.06	0.76	0.06
écrit puis oral	12	0.73	0.06	0.74	0.05
Total	24	0.76	0.07	0.75	0.05
<b>Reprises Mots uniques des énoncés de commande</b>					
oral puis écrit	12	0.71	0.03	0.77	0.07
écrit puis oral	12	0.83	0.08	0.76	0.06
Total	24	0.77	0.08	0.77	0.06

Figure 3 : Tableau des moyennes (m) et écarts-types (E.T.)

## DISCUSSION - CONCLUSION

L'analyse des données a conforté l'hypothèse que les comportements collaboratifs sont présents en DHM. De plus, elle a montré que les indices proposés, Enoncés collaboratifs et Reprises des énoncés de commande, sont à même de rendre compte de différences en fonction des modes d'interaction.

Le premier indice Enoncés Collaboratifs a permis de mettre en évidence que les dialogues avec le service *ArtimisVoyage* étaient plus collaboratifs à l'oral qu'à l'écrit. Plusieurs hypothèses explicatives peuvent être avancées pour rendre compte de ce résultat. Par exemple, bien que la littérature ait montré que les informations verbales vocales soient mieux mémorisées que les informations verbales écrites, le maintien et le traitement des informations à l'oral a semblé être plus difficile. En effet, la permanence de l'écrit peut expliquer que l'utilisateur qui voit à l'écran les critères émis se centre sur la tâche. A l'inverse, l'évanescence de l'oral

peut expliquer que l'utilisateur prenne en compte les demandes du système pour n'oublier aucun critère nécessaire à la recherche. Une des caractéristiques du service était le nombre important de critères à fournir (cinq). Enfin, cette interprétation est cohérente avec les travaux de Clark et Brennan sur les contraintes imposées par les médias et les coûts qu'elles entraînent [12].

L'analyse sur les Reprises des énoncés de commande a montré qu'à l'oral les utilisateurs avaient tendance à partager un lexique et une structure des énoncés commune avec le système. Ce résultat est cohérent avec les travaux de Brennan [4, 5, 6], qui avait, par exemple, montré qu'après quelques interactions avec un partenaire électronique (en Magicien d'Oz) les utilisateurs tendaient à adopter les mêmes modèles syntaxiques que leurs partenaires [4]. Cet entraînement syntaxique peut être attribué à des processus automatiques, en référence aux travaux de Bock [3] sur l'amorçage syntaxique (cf. aussi [19]). Brennan [7] a proposé qu'il soit considéré comme un processus stratégique du discours dans lequel les gens évitent les problèmes en faisant savoir à leur partenaire qu'ils croient qu'ils parlent tous deux de la même chose.

Par ailleurs, nos résultats illustrent aussi le fait que les techniques utilisées pour construire le terrain commun peuvent être différentes en fonction des médias [12]. Cela expliquerait que certains indices de grounding soient moins présents dans certaines modalités [12]. Ainsi, Brennan et Ohaeri [9] ont montré qu'il était plus coûteux de développer les formules de politesse à l'écrit qu'à l'oral. En outre, il serait intéressant de réitérer l'expérimentation présentée ici avec un service différent pour valider les critères d'analyse proposés sur d'autres systèmes mais aussi sur des comparaisons dans des situations de dialogues finalisés homme-homme.

Cette recherche s'inscrit dans un travail visant à fournir des éléments d'analyse et de recommandation sur la collaboration pour les concepteurs de systèmes de DHM. L'amélioration de la collaboration en DHM est un enjeu majeur. Pour atteindre l'utilisabilité optimale des services de dialogue grand public, les systèmes devront tendre vers la prise en charge de comportements typiquement collaboratifs, comme la négociation. Les indices présentés constituent les premières briques dans l'élaboration d'une méthodologie d'analyse de la collaboration. En se centrant sur les comportements de l'utilisateur, elle aura pour objectif d'évaluer en quoi la situation (en fonction de ses caractéristiques, cf. [16]) est propice à la collaboration.

## BIBLIOGRAPHIE

1. Allwood, J., Traum, D. and Jokinen, K. *Cooperation, dialogue and ethics*. Human-Computer Studies, Vol. 53, 2000, pp. 871-914.
2. Amalberti, R., Carbonell, N. and Falzon, P. *User representations of computer systems in human-computer speech interaction*. International Journal of Man-Machine Studies, Vol. 38, 1993, pp. 547-566.
3. Bock, J.K. *Syntactic persistence in language production*. Cognitive Psychology, Vol. 18, 1986, pp. 355-387.
4. Brennan, S.E. *Conversation with and through Computers*. User Modeling and User-Adapted Interaction, Vol. 1, 1991, pp. 67-86.
5. Brennan, S.E. *Lexical Entrainment in Spontaneous Dialog*. Paper presented at International Symposium on Spoken Dialogue, Philadelphia, PA, 1996.
6. Brennan, S.E. *The Grounding Problem in Conversations With and Through Computers*. In S.R. Fussell, R.J. Kreuz (Eds.), *Social and cognitive psychological approaches to interpersonal communication*. Hillsdale, NJ, Lawrence Erlbaum, 1998, pp. 201-225.
7. Brennan, S.E. *The Vocabulary Problem in Spoken Dialogue Systems*. In S. Luperfoy (Ed.), *Automated Spoken Dialog Systems*. Cambridge, MA: MIT Press, 1998.
8. Brennan, S.E., and Clark, H.H. *Lexical choice and conceptual pacts in conversation*. Journal of Experimental Psychology: Learning, Memory, and Cognition, Vol. 22, 1996, pp. 1482-1493.
9. Brennan, S.E. and Ohaeri, J.O. *Why do electronic conversations seem less polite? The costs and benefits of hedging*. Paper presented at Conference on Work Activities, Coordination, and Collaboration (San Francisco), 1999.
10. Clark, H.H. *Using Language*. Cambridge: Cambridge University Press, 1996.
11. Clark, H.H. *Speaking in time*. Speech Communication, Vol. 36, 2002, pp. 5-13.
12. Clark, H.H. and Brennan, S.E. *Grounding in Communication*. In L.B. Resnick, J. Levine, S.D. Teasley (Eds.), *Perspectives on Socially Shared Cognition*. Washington D.C.: APA, 1991, pp. 127-149.
13. Clark, H.H. and Schaefer, E.F. *Collaborating on Contributions to Conversations*. Language and Cognitive Processes, Vol. 2, 1987, pp. 1-23.
14. Clark, H.H. and Schaefer, E.F. *Contributing to Discourse*. Cognitive Science, Vol. 13, 1989, pp. 259-294.
15. Clark, H.H. and Wilkes-Gibbs, D. *Referring as a Collaborative Process*. Cognition, Vol. 22, 1986, pp. 1-39.

16. Dillenbourg, P. *What do you mean by collaborative learning?* In P. Dillenbourg (Ed.), *Collaborative-learning: Cognitive and Computational Approaches*. Oxford: Elsevier Science, 1999, pp. 1-19.
17. Falzon, P. *Ergonomie cognitive du dialogue*. Grenoble: Presse Universitaire, 1989.
18. Falzon, P., Amalberti, R. and Carbonell, N. *Dialogue Control Strategies in Oral Communication*. In K. Hopper, I.A. Newman (Eds.), *Foundation for Human-Computer communication*. North-Holland: Elsevier Science Publishers, 1986.
19. Garrod, S. and Pickering, M.J. *Why is conversation so easy?* *Trends in Cognitive Sciences*, Vol. 8, No. 1, 2004, pp. 8-11.
20. Johnstone, A., Berry, U., NGuyen, T. and Asper, A. *There was a long pause: influencing turn-taking behaviour in human-human and human-computer spoken dialogues*. *International Journal of Human-Computer Studies*, Vol. 41, 1994, pp. 363-411.
21. McCarthy, J.C. and Monk, A.F. *Measuring the quality of computer-mediated communication*. *Behaviour & Information Technology*, Vol. 13, No. 5, 1994, pp. 311-319.
22. Metzging, C. and Brennan, S.E. *When conceptual pacts are broken: Partner-specific effects in the comprehension of referring expressions*. *Journal of Memory and Language*, Vol. 49, 2003, pp. 201-213.
23. Minker, W., and Néel, F. *Développement des technologies vocales*. *Le Travail Humain*, Vol. 65, No. 3, 2002, pp. 261-287.
24. Sadek, D. *Le dialogue homme-machine : de l'ergonomie des interfaces à l'agent intelligent dialoguant*. *Nouvelles Interfaces Homme-Machine*, Vol. 18, 1996, pp. 227-321.
25. Sadek, D., Bretier, P., & Panaget, F. (1997). *Artimis: Natural dialogue meets rational agency*. Paper presented at the Proceedings of the 15th International Joint Conference on Artificial Intelligence (IJCAI'97), Nagoya, Japan.
26. Stent, A., Walker, M., Whittaker, S., and Maloor, P. *User-Tailored Generation for spoken dialogue: An Experiment*. Paper presented at the International Conference on Spoken Language Processing, 2002.
27. Traum, D.R. *Computational Models of Grounding in Collaborative Systems*. *AAAI Fall Symposium on Psychological Models of Communication*, 1999.