

Études expérimentales sur le modèle de l'interlocuteur en dialogue personne-machine

Ludovic Le Bigot - Franck Panaget

France Télécom
Recherche et Développement
2, avenue Pierre Marzin
22300 LANNION Cedex

{ludovic.lebigot ; franck.panaget}
@francetelecom.com

Éric Jamet

Université de Rennes II – CRPCC
6, avenue Gaston Berger
25043 Rennes Cedex
eric.jamet@uhb.fr

Jean-François Rouet

Université de Poitiers – LaCo – CNRS
99, avenue du Recteur Pineau
86000 POITIERS
jean-francois.rouet@univ-poitiers.fr

RESUME

Ce papier présente les résultats de deux études expérimentales contribuant à l'amélioration d'un modèle utilisateur. L'objectif était d'étudier la transformation de la représentation de l'interlocuteur chez l'utilisateur d'un système de dialogue en langage naturel. Plus précisément, le but était de mettre en évidence l'influence du modèle de l'interlocuteur chez les participants sur les nombres de mots et d'informations pertinentes au premier tour de parole. Les résultats ont montré qu'au fur et à mesure des interactions, un utilisateur fournissait de plus en plus d'informations pertinentes pour la recherche au premier tour de parole tout en gardant un nombre de mots par énoncé constant. Ce genre de résultats devrait permettre aux concepteurs de systèmes de dialogue de constituer un outil de détection du niveau de familiarisation de l'utilisateur afin d'améliorer le confort d'utilisation au cours du dialogue.

MOTS CLES : modèle utilisateur, modèle de l'interlocuteur, dialogue, langage naturel, modalité

ABSTRACT

This paper shows the results of two experimental studies contributing to the improvement of a user model. This work aimed at studying the evolution of user's representation when using a dialogue system. More precisely, the purpose was to study the influence of the interlocutor model on the number of words and the number of relevant information in the first dialogue turn. Results has shown that, as experience progress, users produced more and more relevant information for the search in first turn request while keeping a constant word number per utterance. These results should enable dialogue system designers to create a tool that will detect the familiarization level of users. The satisfaction of users will thus be improved during the dialogue.

KEYWORDS: user model, interlocutor's model, dialogue, natural language, modality

INTRODUCTION

Positionnement du problème

Les modèles utilisateur sont devenus dans les interactions personne-machine, et à plus forte raison dans les dialogues personne-machine, un élément critique pour le développement de systèmes et de services utilisables [6]. Le modèle utilisateur constitue une connaissance a priori ou construite en dynamique sur l'utilisateur. Il est mis en œuvre dans des systèmes de dialogue personne-machine dans le but d'adapter l'interaction à un utilisateur particulier ou à un ensemble d'utilisateurs (e.g., [12]).

Dans ce contexte, l'idée était, à partir des recherches sur la représentation des capacités du système par un utilisateur (et de leur influence sur l'interaction), de fournir aux développeurs des éléments pour évaluer le niveau de familiarité de l'utilisateur avec le système.

Le modèle de l'interlocuteur pour enrichir le modèle utilisateur

Des recherches en cognition sociale [3] [5] ont mis en évidence l'influence des connaissances et des croyances initiales des locuteurs sur la construction des messages. Par exemple, Isaacs et Clark [3] ont montré que des locuteurs experts sur New York adaptaient rapidement leurs références en fonction des points de repères connus ou non de leurs destinataires. Fussell et Krauss ([5], exp1) ont montré que les jugements d'individus sur l'expertise des autres étaient corrélés avec leur propre familiarité. Dans une seconde expérience, les auteurs ont montré que les individus utilisaient leurs croyances antérieures pour évaluer les connaissances de leur destinataire. De fait, la construction de leurs messages, même dans des contextes communicatifs où le feedback du locuteur était disponible, était influencée par leurs connaissances réelles ou supposées sur leur(s) interlocuteur(s). Des résultats proches ont été trouvés en dialogue personne-machine. Le paradigme utilisé était un *magicien d'Oz*. Il consiste à faire croire à l'individu

qu'il interagit avec un système alors qu'en réalité un humain lui envoie les messages. Par exemple, une étude [1] s'appuyant sur ce type de paradigme a montré que, dans une tâche orientée par un but, les individus modifiaient la formulation de leur message en fonction du type d'interlocuteur : système ou opérateur humain.

Amalberti et al. ont défini le modèle de l'interlocuteur comme basé « *sur la représentation des compétences techniques et linguistiques de l'interlocuteur, laquelle est en partie élaborée lors du processus de dialogue* » (p.551). En fonction du type d'interlocuteur (humain ou système), un tel modèle influence le comportement verbal des individus. Les auteurs ont trouvé que le nombre d'informations pertinentes (i.e., les informations utiles) pour effectuer une demande au premier tour de parole était moindre dans le groupe système que dans le groupe opérateur humain. Les connaissances erronées des utilisateurs sur les capacités de compréhension et de résolution de problèmes du système les ont amenés à fournir moins d'informations pertinentes pour leur première requête, lorsqu'ils interagissaient avec le système plutôt qu'avec un opérateur humain. Après une série de recherches avec le système, ce nombre d'informations avait tendance à augmenter. D'autres auteurs [7] ont obtenu des résultats voisins. Ces derniers ont permis de mettre en évidence une augmentation du nombre de termes pertinents pour la recherche par énoncé tout en observant des énoncés de plus en plus concis (suppression des termes non essentiels tels que les articles ou les formules de politesse). Dans le même sens lors d'un dialogue personne-machine écrit, une recherche [2] a montré que le nombre d'informations pertinentes pour réaliser la tâche était moins important pour le groupe système lors des premiers tours de parole, puis augmentait dans la seconde partie du dialogue. Des résultats inverses ont été trouvés pour le groupe interagissant avec un humain. Des observations complémentaires sur une comparaison écrit (taper/lire) / oral (parler/écouter) ont été rapportées [13]. Au fur et à mesure des dialogues, la longueur moyenne des énoncés avait tendance à rester stable à l'oral, tandis qu'elle avait tendance à augmenter au clavier.

En résumé, deux résultats importants semblent se dégager des recherches. (1) Le nombre d'informations pertinentes par énoncé pour une recherche d'informations en langage naturel semble augmenter au fur et à mesure des interactions. Ceci est particulièrement vrai au premier tour de parole. (2) La longueur des énoncés semble rester stable, voire augmenter, au fur et à mesure des interactions, au moins à l'oral. Ceci est particulièrement vrai au premier tour de parole.

A partir de ces constats, il est possible de faire les prédictions suivantes. A l'oral, au fur et à mesure des interactions avec un système réel de dialogue en langue

naturelle, (a) le nombre d'informations pertinentes pour une recherche au premier tour de parole devrait augmenter, et (b) le nombre de mots au premier tour de parole devrait rester stable. Pour une interaction à l'écrit, seule la prédiction (a) peut être formulée.

Afin de tester ces hypothèses et d'intégrer éventuellement les résultats au modèle utilisateur, deux expérimentations ont été menées avec un système réel de dialogue personne-machine. L'objectif était de déterminer si pour une application particulière (recherche d'informations), les données de la littérature étaient confirmées avec un système réel.

EXPERIMENTATIONS

Méthode générale

Deux expérimentations séparées ont été menées avec le même système (PlanResto). La principale différence entre les expériences était les connaissances initiales en informatique des participants¹. Dans la première expérience, la population était plutôt familiarisée avec l'informatique alors que, dans la seconde, elle était considérée comme normale (individus tout-venant). Il y a deux raisons essentielles pour distinguer les individus en fonction de leurs connaissances initiales. En premier lieu, Falzon [4] a évoqué l'importance de ces connaissances lors de l'utilisation d'un système de dialogue. Les comportements et les représentations des utilisateurs sont dépendants de leur niveau de familiarité avec un système et avec l'informatique, ou les deux. Par ailleurs, la vitesse de frappe au clavier, et par conséquent la familiarisation avec un environnement informatisé, devraient être plus importantes chez des informaticiens. Que ce soit à l'oral ou à l'écrit, ce type d'utilisateurs devrait être aussi à l'aise quel que soit le mode d'interaction (oral ou écrit).

Une autre différence était que pour la première expérience, les scénarii ont été donnés sous forme de phrases pré-établies (structure identique) alors que pour la seconde, ils étaient sous forme de listes de critères. Néanmoins, dans les deux expériences l'ordre des critères ainsi que les scénarii étaient contrebalancés. De plus, le nombre de réponses pour la première était fixe (1 réponse possible), alors que pour la seconde, le nombre de réponses variait (1 à 6 réponses possibles). Toutefois, les variables recueillies ne concernaient que le premier tour de parole. Dans les sections suivantes, l'application utilisée pour les expérimentations sera présentée puis successivement les méthodes et les résultats de chacune des expériences.

¹ Pour une discussion sur les connaissances initiales en informatique, voir, [11] dans [4]

Cadre formel du modèle utilisateur

La technologie d'agent rationnel dialoguant ARTIMIS² [8] [9] est basée sur une approche de la communication qui stipule que "*communiquer, c'est agir*" [10]. ARTIMIS est fondé sur la théorie de l'interaction qui unifie, dans un cadre formel de logique modale quantifiée, des attitudes mentales telles que la *croyance* et l'*intention* avec un modèle d'actions (communicatives) telles que *informer* et *demande*. Elle modélise, en particulier, un *équilibre rationnel* entre les attitudes mentales d'un agent et ses plans d'actions. Le dialogue est alors modélisé comme l'observation et la planification d'actions communicatives. Dans ARTIMIS, cette théorie est mise en œuvre par un moteur d'inférences : les capacités d'interaction de l'agent dialoguant résultent ainsi d'un processus de raisonnement explicite.

Concernant le modèle utilisateur, la théorie de l'interaction permet d'exprimer finement les croyances d'un agent (1) sur les attitudes mentales (croyance, intention, etc.) d'un autre agent (attitudes actuelles mais aussi passées) et (2) sur les actions réalisées par ou pour cet autre agent. La théorie de l'interaction permet aussi de représenter l'aspect dynamique (au cours du temps) du modèle de l'utilisateur. Ce modèle de l'utilisateur a un pouvoir expressif très précis, mais il manque d'informations sur le "*profil*" de l'utilisateur afin d'adapter les réactions du système.

Application PlanResto

L'application *PlanResto*, est un prototype de service Grand Public à base de dialogue intelligent en langage naturel oral ou écrit pour la recherche de restaurants à Paris (intra-muros). Il est basé sur la technologie ARTIMIS.

L'utilisateur pouvait effectuer une recherche avec trois critères : l'emplacement du restaurant (stations de métro, quartiers, arrondissements, principaux monuments), la gamme de prix et la spécialité culinaire. Le système lui proposait en retour des restaurants correspondant ou s'approchant le mieux de sa requête (cf. des exemples de dialogue, figures 1 et 2).

Deux phases peuvent décrire une recherche typique avec PlanResto. (1) Une phase d'émission des critères, i.e. l'utilisateur formule une demande plus ou moins précise en langue naturelle et le système l'aide en lui demandant de préciser ses critères (s'il le désire, l'utilisateur peut fournir plusieurs critères en un tour de parole). (2) Une phase d'affinage, i.e. l'utilisateur a émis ses critères et suivant les réponses du système, il consulte et parcourt

les solutions. Il peut aussi demander des informations ou des précisions sur les solutions. A tout moment, l'utilisateur peut basculer entre les phases 1 et 2.

S: Bienvenue sur PlanResto. Que puis-je pour vous ?
U: euh, je cherche un restaurant.
S: Vous recherchez un restaurant. J'en ai trouvé plus de 10. Vous pouvez donner plus de précision, par exemple, le lieu, consulter les solutions ou effectuer une nouvelle recherche.
U: près du métro Voltaire
S: Vous recherchez un restaurant près du métro Voltaire. J'en ai trouvé plus de 10...

Figure 1 : Exemple de dialogue avec l'application PlanResto

Le service PlanResto fonctionne en mode vocal par téléphone (moyennant l'intégration des composants de pilotage de la carte téléphonique, de reconnaissance et synthèse de parole au système de dialogue ARTIMIS) mais également en mode écrit via une interface Web.

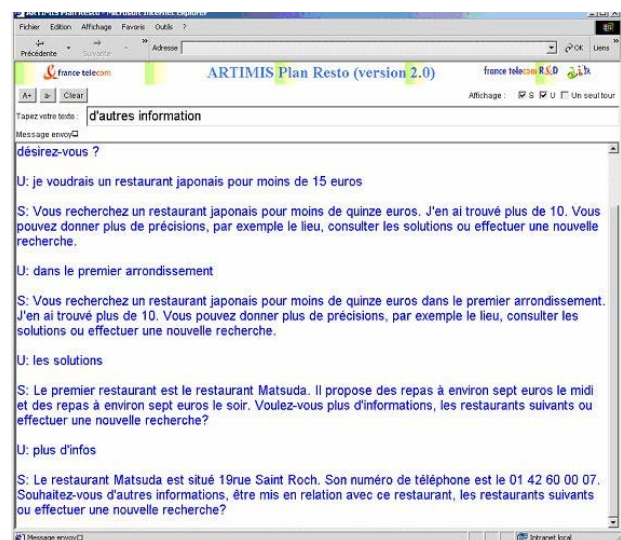


Figure 2 : Interface de dialogue écrit avec PlanResto

EXPERIENCE 1 : POPULATION D'INFORMATIENS Méthode

Participants. Les participants vingt-quatre participants (16 hommes, 8 femmes) ont été répartis de manière équivalente dans deux groupes expérimentaux. La répartition a été faite en fonction de l'âge, du niveau d'étude, des spécialisations professionnelles ou scolaires. L'âge moyen était de 25.8 ans (écart-type, $E.T.=3.9$). Le niveau d'étude était élevé (moyenne, $M=5.2$ années d'étude après le bac, $E.T.=1.59$). Quatorze des participants avaient une formation directement liée à l'informatique, les autres pouvaient être considérés comme fortement familiarisés avec la technologie (usage régulier du courrier électronique, Internet et ordinateur en général). Aucun des participants n'employait régulièrement de services interactifs en

² ARTIMIS est l'acronyme d'"Agent Rationnel fondé sur une Théorie de l'Interaction mise en œuvre dans un Moteur d'Inférence Syntaxique" Voir http://www.rd.francetelecom.fr/fr/galerie/navig_dialogue.htm <http://www.rd.francetelecom.com/fr/technologies/ddm200210/dossier.php>

ligne ou par téléphone. Aucun n'avait utilisé l'application PlanResto auparavant.

Matériel. Afin de mettre l'utilisateur en situation de test, des tâches de recherche d'informations sous forme de scénarii ont été créées. Deux ou trois critères de recherche étaient fournis au participant (lieu, spécialité, prix) pour effectuer sa recherche, il n'y avait qu'une seule réponse pour chaque scénario. En manipulant le nombre et l'ordre des critères, douze questions simples ont été produites. Il existait un scénario par ordre possible et l'ordre des scénarii a été contrebalancé.

Procédure et mesures. Les sujets ont été accueillis individuellement et le service PlanResto leur a été présenté. Ils ont reçu une consigne de résolution de problème : trouver des restaurants à l'aide du service. Les participants n'ont pas été familiarisés avec le service. La seule indication donnée était que le système utilisait le dialogue en langue naturelle. Aucune autre précision n'était apportée. L'expérimentation n'excédait pas 40 minutes.

Le nombre d'informations pertinentes pour la recherche, i.e. le nombre de critères de recherche appropriés fournis par l'utilisateur et pris en compte par le système (de 0 à 3), et le nombre de mots au premier tour de parole ont été recueillis.

Résultats et discussion

Les données ont fait l'objet d'une analyse de la variance pour plan mixte. Les facteurs étaient le mode d'interaction (vocal, écrit) en intergroupe et la position sérielle du dialogue (1 à 12) en intragroupe. Les tests de sphéricité de Mauchly pour le facteur intragroupe étant significatifs ($p < .01$), les degrés de liberté pour les analyses intragroupes ont été ajustés avec un ϵ de Huynh-Feldt (.74 pour le nombre d'informations pertinentes, et .80 pour le nombre de mots). Les données sont représentées sur la figure 3.

Nombre d'informations pertinentes au premier tour de parole. L'analyse a révélé un effet du mode d'interaction ($F(1,22)=10.30, p < .01$) et de la position sérielle du dialogue ($F(8,181)=4.86, p < .05$). Le nombre d'informations pertinentes au premier tour de parole était plus important en mode écrit qu'en mode oral. De plus, il avait tendance à augmenter de manière continue quelle que soit la modalité. L'analyse des tendances a montré une composante linéaire aussi bien en mode oral ($F(1,22)=4.51, p < .05$) qu'en mode écrit ($F(1,22)=5.37, p < .05$). Aucune autre comparaison n'a atteint la significativité statistique.

Nombre de mots au premier tour de parole. L'analyse n'a pas révélé d'effet du mode d'interaction, de la position sérielle du dialogue, ni d'interaction entre les deux facteurs ($F_s < 1, p > .1$).

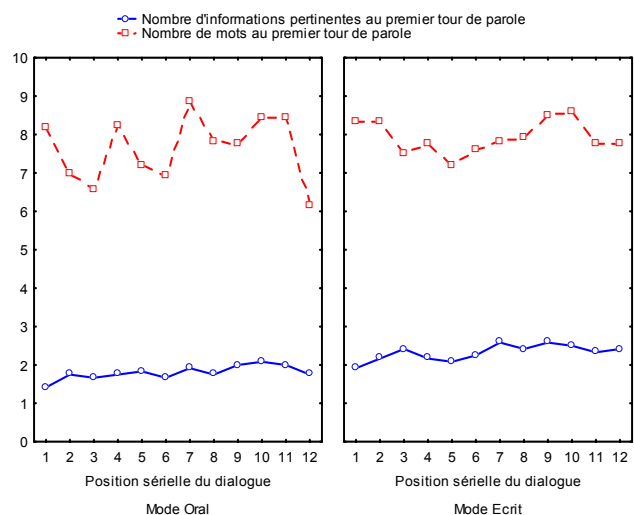


Figure 3 : Effets du mode d'interaction et de la position sérielle du dialogue sur le nombre d'informations et de mots au premier tour de parole chez des individus familiarisés avec l'informatique

Chez des individus familiarisés avec l'informatique, le nombre d'informations pertinentes pour une recherche de restaurants a été plus important en mode écrit qu'en mode oral. De plus, il avait tendance à augmenter. Par contre, quelle que soit la modalité, le nombre de mots au premier tour de parole n'était pas significativement différent suivant la position du dialogue. Ces résultats ont conforté [1] et [13].

EXPERIENCE 2 : POPULATION TOUT-VENANT

Méthode

Participants. Les vingt-quatre participants (12 hommes, 12 femmes) ont été répartis de manière équivalente dans deux groupes expérimentaux. La répartition a été faite en fonction de l'âge, du niveau d'étude, des spécialisations professionnelles ou scolaires. L'âge moyen était de 31.3 ans ($E.T.=8.0$). Le niveau d'étude était de 1.8 années d'étude après le bac ($E.T.=2.51$). Seuls quatre des participants avaient une formation ou un métier directement lié à l'informatique (dont deux linguistes), onze participants étaient peu familiarisés avec la technologie, les treize autres étaient plutôt familiarisés (usage du courrier électronique, Internet et ordinateur en général). Un seul des participants employait régulièrement des services interactifs en ligne ou par téléphone. Aucun n'avait utilisé l'application PlanResto auparavant.

Matériel, procédure et mesures Le même prototype de service Grand Public a été utilisé que dans l'expérience 1. Sur le même principe, douze scénarii de dialogue ont été conçus. Les deux ou trois critères de recherche étaient présentés sous forme de liste. Le nombre de solutions possibles pour chaque scénario variait de 1 à 6. L'ordre des scénarii, le nombre de critères ainsi que le nombre de solutions ont été contrebalancés. La

procédure et les mesures recueillies étaient analogues à l'expérience 1.

Résultats et discussion

Les données ont fait l'objet d'une analyse de la variance pour plan mixte. Les facteurs étaient le mode d'interaction (oral, écrit) en intergroupe et la position sérielle du dialogue (1 à 12) en intragroupe. Les données sont représentées sur la figure 4.

Nombre d'informations pertinentes au premier tour de parole. L'analyse a révélé un effet de la position sérielle du dialogue ($F(11,242)=4.86, p<.0001$) mais pas du mode d'interaction ($F(1,22)=3.08, p>.08$). Le nombre d'informations pertinentes au premier tour de parole avait tendance à augmenter de façon continue quel que soit le mode d'interaction. L'analyse des tendances a montré une composante linéaire aussi bien en mode oral ($F(1,22)=16.44, p<.001$) qu'en mode écrit ($F(1,22)=7.55, p<.05$). Aucune autre comparaison n'a atteint la significativité statistique.

Nombre de mots au premier tour de parole. L'analyse n'a pas révélé d'effet du mode d'interaction, de la position sérielle du dialogue, ni d'interaction entre les deux facteurs ($F_s<1.05, p>.1$).

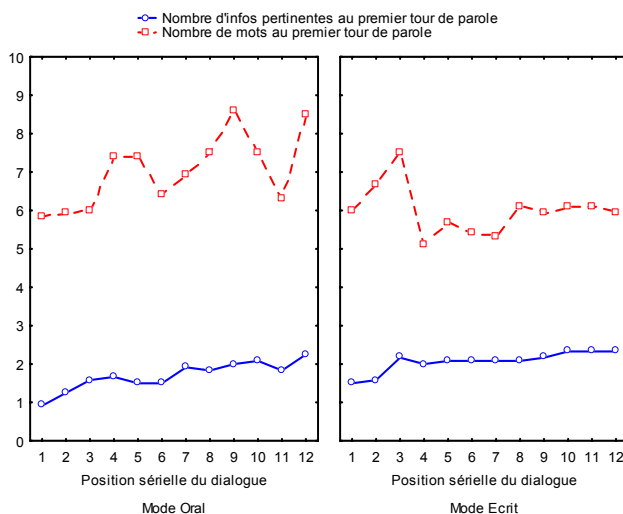


Figure 4 : Effets du mode d'interaction et de la position sérielle du dialogue sur le nombre d'informations et de mots au premier tour de parole chez des individus tout-venant

Chez des individus tout-venant, le nombre d'informations pertinentes pour une recherche de restaurant avait tendance à augmenter quel que soit le mode d'interaction. Par contre, quelle que soit la modalité, le nombre de mots au premier tour de parole n'était pas significativement différent suivant la position du dialogue. Comme pour l'expérience 1, ces résultats ont conforté [1] et [13].

DISCUSSION GENERALE ET CONCLUSION

L'analyse des données a conforté les résultats de [1] et [13] lors de l'utilisation d'un système de dialogue réel pour une recherche d'informations en langue naturelle. Quel que soit le niveau en informatique des sujets, le nombre d'informations pertinentes pour la recherche a augmenté à l'oral comme à l'écrit. Néanmoins, les individus familiarisés avec l'informatique fournissaient plus d'informations en mode écrit. Enfin, que ce soit à l'oral ou à l'écrit, le nombre de mots au premier tour de parole n'était pas significativement différent (dispersion importante). En conséquence, fournir des informations au système sur le savoir-faire de l'utilisateur devrait permettre dans certains cas de procurer des informations complémentaires (e.g. des messages d'aide) pour des individus peu familiarisés ou de supprimer des informations non pertinentes pour un individu familiarisé. Plus généralement, les informations sur le profil des utilisateurs devraient permettre au système d'adapter ses réactions.

Plusieurs pistes peuvent être envisagées pour formaliser les résultats obtenus dans le cadre de la technologie ARTIMIS. Par exemple, plusieurs schémas de propositions devraient permettre à un agent (1) de se représenter le niveau d'expertise de son interlocuteur et (2) de fournir, le cas échéant, des conseils d'utilisation. Par exemple, il est possible de faire une estimation du rapport entre le nombre d'informations pertinentes et le nombre de mots au premier tour de parole pour en déduire le degré de familiarisation avec le système des utilisateurs. Dans le contexte d'un service de recherche de restaurants en langue naturelle avec trois informations pertinentes maximum pour la recherche, deux niveaux peuvent être distingués :

- (1) Un utilisateur peu expérimenté fournit au premier tour de parole peu d'informations pertinentes en un nombre de mots relativement élevé.
- (2) Un utilisateur expérimenté fournit au premier tour de parole un nombre d'informations pertinentes important en un nombre de mots relativement limité.

Mais la mise en évidence de la différence entre mode écrit et mode oral chez des individus familiarisés avec l'informatique incite à relativiser les résultats. Les deux niveaux précédemment cités ne devraient être exploités qu'en mode oral.

La détection du niveau de familiarité de l'utilisateur avec un service est un objectif important pour les concepteurs d'applications informatiques. Elle permettrait d'adapter le service aux attentes et aux capacités de l'utilisateur. La constitution d'un modèle utilisateur fiable et dynamique est une des solutions pour atteindre ce but. Cette étude est la première partie d'un travail sur la détection des profils utilisateurs. La récupération et l'exploitation

d'indicateurs comportementaux devraient permettre au(x) système(s) de s'adapter de manière dynamique aux utilisateurs. Il reste maintenant à savoir si l'amélioration du modèle utilisateur par la suppression d'énoncés système non pertinents pour un utilisateur régulier augmente l'efficacité et la satisfaction. Dans le même sens, est-ce que fournir des conseils en dynamique lors d'une interaction dialogique améliorerait la satisfaction et l'efficacité de l'utilisateur sans alourdir sa charge de travail mentale ?

Enfin, plusieurs limites à cette étude doivent être apportées. Premièrement, bien que l'expérimentation ait été réalisée avec un système réel, elle était relativement contrôlée. Il est possible de se demander si les effets observés ne sont pas dus à une familiarisation avec la tâche (recherche d'informations) plutôt qu'à une familiarisation avec l'utilisation du système. De plus, l'exploitation des données ne sera pertinente que si les tâches (scénarii) sont suffisamment représentatives d'une utilisation réelle du système. Deuxièmement, seule la détection d'une partie des utilisateurs peut être réalisée avec cette méthode. A l'avenir, il sera important de coupler la récupération des indices identifiés avec d'autres indicateurs comportementaux. Une intégration dans un modèle statistique prédictif global pourrait permettre d'affiner le diagnostic du profil utilisateur. Troisièmement, l'utilisation de cette méthode ne permet pas de prédire les comportements des utilisateurs pour un usage sur une période longue (e.g., une utilisation toutes les semaines). Quatrièmement, même si de nombreux tests d'utilisabilité ont été réalisés pour le service *PlanResto*, les données recueillies devront être confrontées à une utilisation réelle après le déploiement du service. Enfin, l'exploitation des résultats n'est valable, pour l'instant, que pour un service avec peu de critères pertinents et sur une tâche de recherche d'informations. Il serait intéressant de reproduire ce genre d'expérimentations sur un système avec plus de critères de recherches et dans un cadre différent.

BIBLIOGRAPHIE

1. Amalberti, R., Carbonnell, N. and Falzon, P. *User representations of computer systems in human-computer speech interaction*. International Journal of Man-Machine Studies, 38, 1993, pp. 547-566.
2. Brennan, S. *Conversation with and through Computers*. User Modeling and User-Adapted Interaction, 1, 1991, pp. 67-86.
3. Isaacs E. and Clark H. *References in conversation between experts and novices*. Journal of Experimental Psychology: General, 2(6) 1987, pp. 26-37.
4. Falzon, P. *Ergonomie Cognitive du Dialogue*. Grenoble: Presses Universitaires de Grenoble, 1989.
5. Fussell, S. R. and Krauss, R.M. *Coordination of Knowledge about in Communication: Effects of Speakers' Assumptions About What Others Know*. Journal of Personality and Social Psychology, 62(3), 1992, pp. 378-391.
6. Litman, D. J. and Pan, S. *Designing and Evaluating an Adaptive Spoken Dialogue System*. User Modeling and User-Adapted Interaction, 12(2-3), 2002, pp. 111-137
7. Richards, M. A. and Underwood, K. M. How should people and computers speak to each other? Paper presented at *the Human-Computer Interaction - INTERACT'84*, London, 1985, 4-7 September.
8. Sadek D., Design considerations on dialogue systems: from theory to technology - the case of Artemis, Proceedings of the *ESCA TR Workshop on Interactive Dialogue for Multimodal Systems (IDS'99)*, invited paper, Germany, 1999.
9. Sadek D., Bretier P. and Panaget F., *Artemis: Natural dialogue meets rational agency*, Proceedings of the *15th International Joint Conference on Artificial Intelligence (IJCAI'97)*, Nagoya, Japan, 1997, pp. 1030-1035.
10. Searle J.R., *Speech acts*, Cambridge University Press, 1969.
11. Thomas, J.C. *A method for studying natural language dialogue* (Rapport RC 5885). Yorktown Heights, New York: IBM Watson Research Center, 1976
12. Whittaker, S., Walker, M. and Moore, J. Fish or Fowl: A Wizard of Oz Evaluation of Dialogue Strategies in the Restaurant Domain. Paper presented at the *Language Resources and Evaluation Conference*, Spain, 2002
13. Zoltan-Ford, E. *How to get people to say and type what computers can understand*. International of Journal Man-Machine Studies, 34, 1991, pp. 527-547.