

Caractérisation d'un corpus de requêtes d'assistance

François BOUCHET

LIMSI-CNRS, Université Paris-Sud XI, BP 133, 91403 Orsay Cedex

bouchet@limsi.fr

Résumé. Afin de concevoir un agent conversationnel logiciel capable d'assister des utilisateurs novices d'applications informatiques, nous avons été amenés à constituer un corpus spécifique de requêtes d'assistance en français, et à étudier ses caractéristiques. Nous montrons ici que les requêtes d'assistance se distinguent nettement de requêtes issues d'autres corpus disponibles dans des domaines proches. Nous mettons également en évidence le fait que ce corpus n'est pas homogène, mais contient au contraire plusieurs activités conversationnelles distinctes, dont l'assistance elle-même. Ces observations nous permettent de discuter de l'opportunité de considérer l'assistance comme un registre particulier de la langue générale.

Abstract. In order to conceive a conversational agent able to assist ordinary people using softwares, we have built up a specific corpus of assistance requests in french, and studied its characteristics. We show here that assistance requests can be clearly distinguished from the ones from other available corpora in related domains. We also show that this corpus isn't homogeneous, but on the contrary reflects various conversational activities, among which the assistance itself. Those observations allow us to discuss about the opportunity to consider assistance as a general language particular registre.

Mots-clés : corpus de requêtes d'assistance, agent conversationnel, activité conversationnelle, actes de dialogue.

Keywords: corpus of assistance requests, conversational agent, conversational activity, speech acts.

1 Introduction

Le développement de l'informatique pour le grand public a entraîné une forte augmentation du nombre d'utilisateurs novices en informatique en contact régulier avec celle-ci. Ces utilisateurs novices n'ont bien souvent ni le temps ni l'envie d'utiliser des manuels papiers ou des FAQ (Foire Aux Questions) de logiciels de plus en plus complexes (en dépit des progrès ergonomiques) dont par ailleurs ils ne maîtrisent pas le vocabulaire spécifique. Des systèmes d'aides contextuelles (ou CHS en anglais, pour Contextual Help Systems (Jansen, 2005)) ont été développées pour mieux s'adapter aux besoins de ces nouveaux utilisateurs, mais ceux-ci semblent toujours préférer faire appel à un ami expert lorsqu'ils souhaitent réaliser une tâche particulière dans une application (Capobianco & Carbonell, 2002).

Parallèlement, les agents conversationnels animés dotés de capacités de dialogue et de raisonnement de niveaux variés (Sadek *et al.*, 1997) développés récemment ont mis en évidence les nom-

breux avantages potentiels d'une présence (même virtuelle) pour faciliter l'interaction homme-machine (Lester *et al.*, 1997).

Pour répondre à ce besoin d'assistance des usagers novices, le projet DAFT développé au LIMSI-CNRS (Sansonnet *et al.*, 2005) se propose donc de développer des Agents Conversationnels Assistants (ACA), capables d'analyser des requêtes en langue naturelle écrite non contrainte provenant d'usagers novices en situation réelle d'utilisation d'applications de complexités diverses (applets simples, pages web, traitement de texte). Pour répondre à ce type de requêtes, le système d'assistance raisonne sur la structure et le fonctionnement des applications à l'aide d'un modèle de celles-ci, construit de manière semi-automatique. Cette méthodologie a pour objectif de fournir une assistance pertinente en contexte à la manière des CHS, avec en plus tous les bénéfices liés à la présence d'agents conversationnels animés.

Afin d'identifier précisément les propriétés et les besoins propres à la Fonction d'Assistance, nous avons été amenés à constituer un corpus de requêtes, auquel on se référera sous le nom de corpus DAFT. Ce corpus illustre les actes de dialogues réalisés par des sujets en situation d'assistance et a été étudié lors d'un stage de M2R (Bouchet, 2006) afin de pouvoir réaliser les spécifications d'un langage de requêtes formelles adapté au domaine d'étude¹.

Dans cet article, nous détaillons dans un premier temps le processus de constitution du corpus employé, en justifiant en particulier la nécessité des choix effectués et en vérifiant que le corpus ainsi obtenu est viable pour répondre à nos besoins. Dans un second temps, nous nous proposons de comparer ce corpus avec d'autres corpus de domaines proches pour en dégager les spécificités. Enfin, nous analysons les différentes activités couvertes par les requêtes recueillies et tentons de les caractériser par des paramètres d'ordre linguistique.

2 Collecte et construction du corpus

2.1 Méthodologie employée

Au moment de l'étude, le corpus DAFT se composait d'environ 5 000 requêtes isolées (cf. §2.3) recueillies entre juin 2004 et juin 2006. Pour le constituer, nous avons eu recours à deux méthodes complémentaires garantissant à la fois l'empirisme et la bonne couverture du corpus :

1. Recueillir des requêtes réelles produites par des utilisateurs placés devant des applications intégrant un ACA de type LEA² (2/3 du corpus final).
2. Utiliser des structures dialogiques génériques issues de classifications de thésaurus anglais (600 structures (Molinsky & Bliss, 1994)) ou bilingues (300 structures (Atkins & Lewis, 1996)). Ces structures ont été adaptées pour les employer dans des requêtes d'assistance formulées dans le contexte des applications mentionnées dans le point précédent, afin d'assurer une certaine homogénéité lexicale du corpus (1/3 du corpus final). Ainsi, la structure "ça te dirait de..." utilisée conjointement avec une phrase recueillie comme "peux-tu jouer à ma place?" donne "ça te dirait de jouer à ma place?".

¹Une première version moins complète de cette étude a été présentée lors de l'atelier WACA'02, n'ayant pas donné lieu à une parution d'actes.

²LIMSI Embodied Agent, développé par J-C. Martin dans le cadre du projet NICE (Buisine & Martin, 2005)

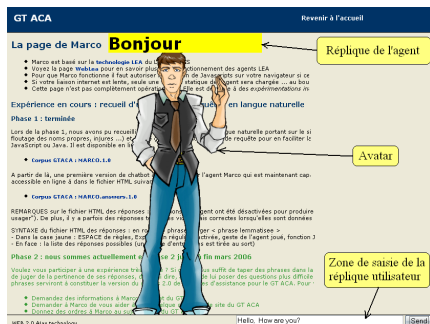


FIG. 1 – L'agent WebLea (Marco) placé sur une page du site web du GT ACA

L'utilisation de la seconde méthode nous permet de compenser la difficulté à obtenir un corpus de taille supérieure avec le panel réduit de sujets dont nous disposons³. Il y a recouvrement partiel avec les phrases recueillies par la première méthode, mais ce choix, discutable dans l'absolu, doit être mis en relation avec l'objectif de ce corpus qui est de fournir une base pour la constitution d'un agent rationnel : la couverture prime donc sur l'exactitude de la fréquence des phénomènes linguistiques. Le corpus ainsi construit sature donc mieux le domaine d'étude, même si cela induit fatalement un biais en sur-représentant des phénomènes linguistiques rares (les structures employées ne seraient apparues "naturellement" que sur un corpus plus grand).

2.2 Recueil des requêtes d'assistance

Les phrases recueillies l'ont été au sein de deux types d'applications :

1. deux applications de type Java (Le Guern, 2004) : un simple compteur temps réel (thread Java) dont l'utilisateur contrôle le démarrage et la vitesse, et un jeu de tours de Hanoi fonctionnant de manière modale (ie n'évoluant que si l'utilisateur agit).
2. deux sites web : une version active du site du groupe AMI du LIMSI permettant l'édition de contenu, et le site du GT ACA⁴(cf. figure 1) en libre accès sur internet aux utilisateurs effectifs du site.

Dans la mesure où l'agent assistant que nous souhaitons réaliser doit disposer d'une certaine genericité (indépendance par rapport aux applications assistées), ces différentes origines ne posent pas de problème concernant l'homogénéité de notre corpus. L'emploi d'un vocabulaire très spécifique à une application particulière permet parfois de retrouver a posteriori l'origine de certaines requêtes, mais la formulation générale des requêtes n'est elle pas affectée (par exemple, la phrase "comment faire pour arrêter le compteur ?" a une structure globale identique à "comment faire pour s'inscrire au GT ACA ?").

³Environ 50 sujets ayant réalisé une session représentant de 10 à 50 requêtes en environnement contrôlé sur les 3 premières applications décrites, et 30 personnes ayant participé à la campagne "Marco", ouverte sur le web à tous les usagers du GT ACA entre janvier et mars 2006.

⁴Groupe de Travail sur les Agents Conversationnels Animés - <http://www.limsi.fr/aca/>

N°	Phrases du corpus	N°	Phrases du corpus
1	a ppuies sur le bouton quitter	8	j'ai été surpris qu'il manque une fonction d'annulation
2	c lickersur le bouton back	9	ça serait mieux si on pouvait aller directement au début
3	bon, reviens à l apage d'accueil	10	auf viedersen
4	a quoi sert cette fenêtre,	11	ca marche :-)
5	c quoi le GT ACA	12	Quel genre de musique tu aimes ?
6	le bouton "fermer" et le bouton "quitter" ont le même fonctionnement ?	13	bon à rien !
7	je ne v osi aucune page de d emso !!	14	j'aime tes cheveux Léa

TAB. 1 – Exemples de phrases du corpus DAFT avec différentes sortes d'erreurs en gras

2.3 Vue du corpus

La table 1 expose un extrait du corpus DAFT, faisant apparaître certaines de ses caractéristiques :

- beaucoup de phrases sont *bruitées* (expressions orales, fautes d'orthographe, de syntaxe et de grammaire, langage SMS...), et certaines erreurs sont non triviales à traiter avec des outils classiques de TALN.
- il ne se présente *pas* comme une succession de dialogues homme-machine, mais plutôt comme une liste de phrases employées par les usagers (questions, ordres ou remarques à l'agent...). En effet, on constate que dans le cadre de la Fonction d'Assistance, les interactions dialogiques se limitent essentiellement à un seul et unique tour de parole (commande-action, question-réponse...), et peuvent donc être traitées de manière *isolée*⁵.

2.4 De la nécessité du corpus

Bien qu'il n'y ait pas de corpus d'assistance similaires aisément disponibles en français, on est en droit de s'interroger sur la nécessité de constituer un nouveau corpus : est-ce qu'un corpus dans un domaine connexe tel que le dialogue homme-machine orienté tâche n'aurait pas été suffisamment proche pour nous convenir ?

Pour répondre à cette question, et ainsi justifier la nécessité d'un corpus particulier, on peut suivre la méthode de comparaison statistique de corpus présentée dans (Ripoche, 2006) pour comparer le corpus DAFT à plusieurs corpus de dialogues homme-homme orientés tâche par l'étude de leurs profils interactionnels.

On appelle profil interactionnel d'un corpus une représentation sous forme d'histogrammes de la répartition des différents actes de dialogue (au sens de (Searle, 1969)) au sein de celui-ci (cf. fig. 2). Les trois corpus de référence choisis pour effectuer cette comparaison sont Switchboard (Jurafsky *et al.*, 1998) (200 000 énoncés de conversations téléphoniques orientées tâche annotés manuellement), MapTask (Carletta *et al.*, 1996) (128 dialogues visant à reconstruire une carte par placement de points de repère) et Bugzilla (Ripoche, 2006) (1 200 000 commentaires issus de 128 000 rapports de défauts établis lors du développement de la suite logicielle de la Fondation Mozilla).⁶

⁵Ce point essentiel de notre approche nous a été confirmé par N. Carbonell, et sera discuté en détails en 2.5.

⁶Switchboard et MapTask, de part leur nature orale, sont naturellement plus riches en nombre de mots que des corpus écrits (Kelly & Chapanis, 1977), mais la proximité des activités a primé sur cette différence de nature dans notre choix de les employer pour cette comparaison (en l'absence de corpus écrits équivalents connus).

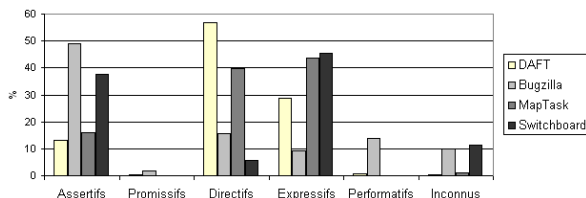


FIG. 2 – Comparaison des profils interactionnels de corpus d'assistance dialogique

Nous avons donc effectué une conversion des taxonomies d'actes de dialogue de ces différents corpus vers une taxonomie commune en 5 actes, tandis qu'un sous-ensemble au 1/5^e du corpus DAFT a été manuellement annoté directement dans cette même taxonomie. Certains actes de dialogues très spécifiques (par exemple les "self-talk" de Switchboard) sont difficiles à convertir ; les résultats ne sont donc pas parfaits et leur interprétation doit être considérée avec certaines précautions. Toutefois, certaines caractéristiques suffisamment nettes distinguent le corpus DAFT des trois autres :

- une présence majoritaire (57 %) de *directifs*, s'expliquant par un nombre élevé d'ordres directs ou de questions à l'agent. Bien qu'orientés tâche, les autres corpus mettent en jeu uniquement des interlocuteurs humains, et il est vraisemblable que le fait de s'adresser à la machine (même via un agent conversationnel) tend à rendre les requêtes plus directes car les usagers supposent que l'agent n'est pas capable des mêmes inférences qu'un être humain.
- un nombre assez faible d'*assertifs* (13 %), l'utilisateur exprimant bien plus son état d'esprit (29 %) par rapport à des faits que ces mêmes faits de manière neutre et "objective" comme c'est le cas par exemple dans le corpus Bugzilla ci-dessus.
- quelques *promissifs* sont présents (1 %) mais marginaux, ce qui s'explique par la nature de la relation agent-utilisateur, car si l'agent est par essence aux ordres de l'utilisateur, ce dernier ne se sent que rarement engagé envers son agent assistant (même pour suivre ses conseils).

Ces divergences entre corpus portant sur un même thème (l'assistance dialogique à une tâche) démontrent que notre hypothèse était justifiée et que notre champ d'étude se distingue suffisamment des champs connexes pour que nous ayons besoin de disposer d'un corpus propre.

2.5 L'assistance : un registre de langue ?

Le fait que la plupart des interactions ne nécessitent qu'un seul tour de parole rapproche finalement davantage notre projet des interfaces d'Interaction en Langue Naturelle (NLI en anglais, comme défini par (Androutsopoulos & Aretoulaki, 2003)) que des systèmes de dialogue à proprement parler, ce qui peut s'expliquer par le fait que le domaine de l'assistance est relativement circonscrit. Il pourrait éventuellement être considéré comme constituant un sous-langage au sens de (Kittredge, 1982) (au même titre que les bulletins météorologiques ou la biologie) ou tout du moins comme un registre particulier de langue (Biber, 1995), dans la mesure où si l'on fait abstraction du clavardage entre l'utilisateur et l'agent qui n'est pas notre centre d'étude principal, on constate globalement que :

- le vocabulaire employé est assez pauvre et fortement lié à l'application assistée (cf. phrases 1-9 de la table 1).
- les classes lexicales les plus fréquentes sont peu nombreuses car essentiellement en rapport avec des éléments de l'interface graphique (boutons, champs texte...) et avec des actions

standards (ajouter, modifier, déplacer...).

- les structures de phrases, particulièrement dans les sous-corpus d’assistance, sont assez prototypiques et très différentes de ce que l’on trouve dans la langue naturelle générale⁷.

3 Catégorisation et caractérisation du corpus DAFT

3.1 Catégorisation des activités conversationnelles

Lors de la phase de recueil du corpus, les sujets humains ont été informés qu’ils devaient réaliser certaines tâches, en faisant appel si nécessaire à un agent (non humain) présent dans l’application pour les assister. Les sujets pouvaient néanmoins agir et s’exprimer de manière non contrainte, et divers comportements ont pu être observés, l’utilisateur se détournant parfois complètement de sa tâche initiale. Finalement, il apparaît donc que de nombreuses phrases recueillies ne relèvent pas vraiment du domaine de l’assistance (cf. table 1). Nous nous sommes par conséquent intéressés à identifier les différentes activités conversationnelles réellement présentes dans le corpus.

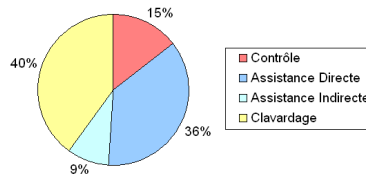


FIG. 3 – Répartition du corpus en sous-corpus par activités

Pour cela, nous avons extrait aléatoirement des phrases du corpus de manière à former deux sous-ensembles de taille égale au dixième de la taille totale du corpus. Dans le premier sous-ensemble, les phrases ont été regroupées manuellement par activités similaires, en s’intéressant notamment au thème des requêtes (application, agent, utilisateur...) et à leur nature (ordre, question, compliment...). On a finalement obtenu ainsi quatre grandes catégories de tailles inégales (cf. fig. 3). Ensuite, connaissant ces quatre activités principales, on a traité le deuxième sous-ensemble en classifiant manuellement les phrases de celui-ci dans une des activités définies précédemment. La répartition ainsi déterminée était très proche de celle trouvée sur le premier sous-ensemble, ce qui nous laisse penser qu’on peut raisonnablement généraliser ce résultat à l’ensemble du corpus (la figure 3 présente les résultats obtenus en faisant la moyenne des deux répartitions trouvées).

On peut par conséquent considérer que notre corpus est divisible en quatre “sous-corpus”, correspondant chacun à des types d’activités distinctes (les phrases données en exemples sont celles de la table 1) :

1. **activité de contrôle** : corpus constitué de *commandes*, afin que l’agent agisse lui-même sur l’application (*phrases 1-3*).
2. **activité d’assistance directe** : regroupant des *demandes d’aide* explicitement formulées comme telles par l’utilisateur (*phrases 4-6*).

⁷Comme l’a montré une autre comparaison effectuée par rapport à un extrait du corpus généraliste Multitag, fourni par P. Paroubek.

Sous-corpus	Contrôle	Assist. directe	Assist. indirecte	Clavardage
Moyenne	5,44	8,01	9,90	6,01
Écart-type	3,36	3,54	3,30	3,62

TAB. 2 – Répartition des phrases par longueur (mots) dans les sous-corpus

3. **activité d'assistance indirecte** : corpus formé d'*opinions* sur l'application qui constituent des demandes d'aide sous-entendues, probablement perceptibles uniquement au niveau pragmatique (*phrases 7-9*).
4. **activité de clavardage** : réunissant le reste des interactions essentiellement centrées sur l'agent ainsi que des expressions métalinguistiques, phatiques⁸ et de backchanneling⁹ (*phrases 10-14*).

L'existence des sous-corpus de contrôle et de clavardage démontre que l'utilisateur attend non seulement d'un ACA qu'il l'aide à utiliser une application, mais aussi qu'il soit capable d'agir lui-même sur celle-ci, ainsi que de répondre à des commentaires annexes indépendants de la tâche à accomplir où l'agent devient lui-même le centre d'intérêt de l'utilisateur (ce phénomène s'expliquant essentiellement par la présence d'une représentation visuelle de l'agent).

3.2 Méthodes de caractérisation des sous-corpus

Les quatre sous-corpus ont été catégorisés précédemment uniquement par une annotation *manuelle*, mais il serait souhaitable de pouvoir automatiser cette classification afin d'analyser spécifiquement les activités propres à chaque sous-corpus. On envisage alors trois méthodes de caractérisation possibles de ceux-ci :

- une étude de la distribution de la longueur des phrases des sous-corpus.
- une étude des profils interactionnels des sous-corpus, tels que définis en 2.4.
- une étude de la sémantique des phrases par analyse de leur retranscription sous forme de requêtes formelles.

3.2.1 Caractérisation par la longueur des phrases

On observe une certaine disparité de longueur des requêtes, les requêtes de contrôle semblant globalement assez courtes comparées aux requêtes d'assistance (cf. table 2), et on peut approximer les répartitions des sous-corpus de contrôle et d'assistance indirecte par une loi normale (test de χ^2 avec un seuil de tolérance de 1%). Néanmoins, les écart-types trop importants ($\sigma \approx 3,5$) disqualifient en pratique cette méthode de classification.

3.2.2 Caractérisation par l'analyse des profils interactionnels

On distingue sur la figure 4 certaines différences de profils interactionnels assez nettes entre les sous-corpus et par rapport au profil générique du corpus DAFT (rappelé en gris foncé), notamment pour distinguer l'assistance directe (avec une forte majorité de directifs et quelques expressifs) de l'assistance indirecte (une majorité d'assertifs et des expressifs). En revanche, les

⁸Pour maintenir le contact communicatif avec l'agent : « pas vrai lea ? », « tu dors ou quoi ? »...

⁹Pour marquer son accord aux propos du locuteur et l'inciter à continuer : « Bon. », « ok ok »...

profils interactionnels des sous-corpus de contrôle et d'assistance directe sont assez similaires. Cette méthode présente donc un certain intérêt mais ne peut être utilisée de manière unique. En outre, automatiser la détection de ces actes de dialogue n'est pas non plus trivial.

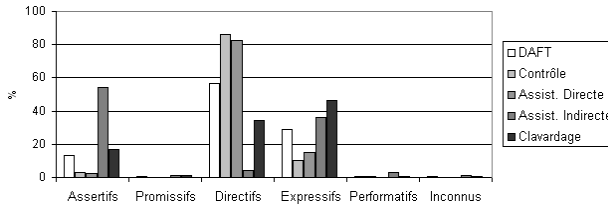


FIG. 4 – Comparaison des profils interactionnels des sous-corpus du corpus DAFT

3.2.3 Caractérisation par analyse des requêtes formalisées

Notre principale motivation pour constituer ce corpus était de pouvoir s'en servir comme base pour la définition d'un langage formel adapté pour exprimer la sémantique des requêtes langagières. La syntaxe de ce langage, présentée en détails dans (Bouchet, 2006), distingue des modalités (la possibilité, l'obligation, le savoir...), des prédicats d'actions (modifier, déplacer, actionner...) et des références (le tableau, le petit bouton rouge...) qui peuvent s'imbriquer entre elles sous la forme :

$$M_1(\dots M_n(c_1 = P_1(c'_1 = R_1, \dots, c'_l = R_l), \dots, c_m = P_m(\dots)) \dots)$$

Expression dans laquelle : $M_1 - M_n$ sont des **modalités**,

$P_1 - P_m$ sont des **prédicats**,

$c_1 - c_m$ & $c'_1 - c'_l$ sont des intitulés de champs typés (par exemple : objet, personne, lieu...),

$R_1 - R_l$ sont des **références** issues de la requête langagière.

Exemple : La phrase 6 de la table 1 pourra ainsi s'écrire sous la forme :

CHECK(NEG(DIFFERENCE(between="le bouton 'fermer'", and="le bouton 'quitter'", is="fonctionnement")))

Si l'on s'intéresse alors au nombre de modalités ou de prédicats présents dans chacun des sous-corpus, on constate des différences assez nettes (cf. figure 5) :

- le contrôle possède un nombre élevé de prédicats par phrase (0.97) et peu de modalités (0.22).
- l'assistance directe contient beaucoup de modalités (2.41) ainsi qu'un nombre moyen de prédicats (0.54).
- l'assistance indirecte diffère peu de l'assistance directe avec 0.59 prédicats par phrase et légèrement moins de modalités (2.22), avec toutefois une part plus forte de phrases à modalité unique (18% contre 5%).
- le clavardage (limité aux phatiques pour cette étude) est extrêmement pauvre en prédicats (0.08) et assez peu de modalités (0.86).

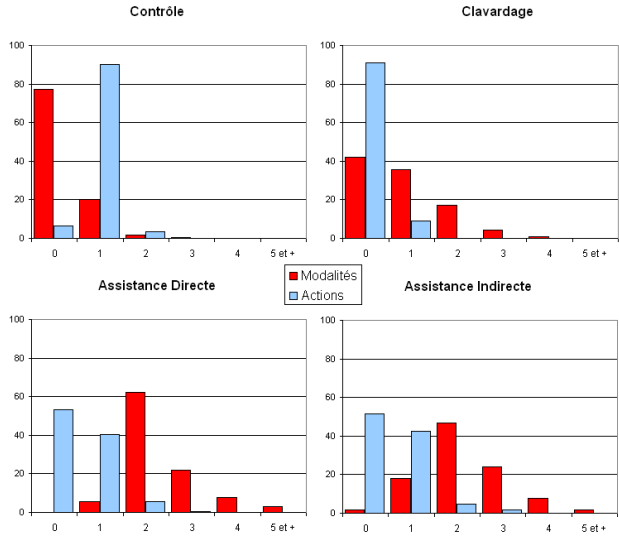


FIG. 5 – Nombre de modalités et prédicats par phrase dans chacun des sous-corpus

3.2.4 Conclusion de l'étude

Parmi les trois méthodes testées, aucune ne permet d'identifier parfaitement l'activité correspondant à une requête. On constate en effet que :

- la longueur des phrases n'apporte pas d'informations discriminantes,
- la comparaison des profils interactionnels permet une discrimination efficace des deux formes d'assistance (directe et indirecte) qui nous intéressent dans le cadre du projet DAFT,
- la comparaison des requêtes mises sous forme formelle permet de distinguer le contrôle, l'assistance et le clavardage.

Pour classifier automatiquement les requêtes dans un des sous-corpus, on pourra donc envisager de combiner ces deux dernières méthodes.

4 Conclusion et perspectives

La Fonction d'Assistance, dans le cadre des CHS, constitue un registre de langue particulier, et se distingue d'activités connexes comme le dialogue homme-homme orienté tâche. Nous avons constitué un corpus composé de requêtes d'assistance recueillies dans le cadre de diverses applications et d'autres requêtes construites en situation, ce qui nous a permis d'avoir un ensemble globalement représentatif de l'assistance. Nous avons montré que ce corpus recouvre en réalité quatre activités distinctes, identifiables par une combinaison de méthodes statistiques classiques et d'une analyse nécessitant la transcription des requêtes dans un langage formel.

Depuis la réalisation de cette étude, le corpus DAFT a été complété, notamment afin de renforcer la proportion de requêtes d'assistance (qui ne représentait ici que 45% du corpus total) en incluant des requêtes relevées face à une application de nature plus complexe (Word). Il compte

ainsi désormais un peu plus de 11 000 phrases, qui nous servent actuellement de base solide pour la construction de la chaîne de traitement des requêtes langagières (analyses grammaticale et sémantique). Un large extrait sous forme brute ainsi que des phrases analysées par notre système sont disponibles en libre accès¹⁰.

Références

- ANDROUTSOPOULOS I. & ARETOULAKI M. (2003). *The Oxford Handbook of Computational Linguistics*, chapter Natural Language Interfaces, p. 629–649. Oxford University Press.
- ATKINS B. & LEWIS H. (1996). *The Collins-Robert French-English Dictionary*, chapter Language in Use. Harper Collins Publishers, 1st edition.
- BIBER D. (1995). *Dimensions of register variation : a cross-linguistic comparison*. Cambridge University Press.
- BOUCHET F. (2006). Conception d'un langage de requêtes pour un agent conversationnel assistant. Master's thesis, Univ. Paris XI.
- BUISINE S. & MARTIN J.-C. (2005). Children's and adults' multimodal interaction with 2d conversational agents. In *Proceedings of the SIGCHI conference on Human factors in computing systems*, p. 1240–1243, Portland, Oregon, USA : ACM Press.
- CAPOBIANCO A. & CARBONELL N. (2002). Conception d'aides en ligne pour le grand public : défis et propositions. In *Actes de ERGO-IA'2002*, p. 309–335.
- CARLETTA J., ISARD A., ISARD S., KOWTKO J., DOHERTY-SNEDDON G. & ANDERSON A. (1996). *HCRC dialogue structure coding manual*. Rapport interne, Univ. of Edinburgh.
- JANSEN B. J. (2005). Seeking and implementing automated assistance during the search process. *Information Processing and Management*, **41**(4), 909–928.
- JURAFSKY D., BATES R., COCCARO N., MARTIN R., METEER M., RIES K., SHRIBERG E., STOLCKE A., TAYLOR P. & VAN ESS-DYKEMA C. (1998). *Switchboard Discourse Language Modeling Project Final Report*. Rapport interne, Center for Speech and Language Processing, Johns Hopkins University, Baltimore, MD, USA.
- KELLY M. J. & CHAPANIS A. (1977). Limited vocabulary natural language dialogue. *International Journal of Man-Machine Studies*, **9**(4), 479–501.
- KITTREDGE R. (1982). Variation and homogeneity of sublanguages. In *Sublanguage : Studies of Language in Restricted Semantic Domains*, p. 107–137. De Gruyter.
- LE GUERN K. (2004). Définition d'une architecture de médiateur pour des agents conversationnels animés. Master's thesis, Univ. Paris XI.
- LESTER J. C., CONVERSE S. A., KAHLER S. H., BARLOW S. T., STONE B. A. & BHOGAL R. S. (1997). The Persona Effect : Affective impact of animated pedagogical agents. In *Proc. of CHI'97*, p. 359–366, New York, NY, USA : ACM Press.
- MOLINSKY S. J. & BLISS B. (1994). *Inventory of functions and conversation strategies*, In *Communicator : The Comprehensive course in functional English*, p. 177–187. Prentice Hall.
- RIPOCHE G. (2006). *Sur les traces de Bugzilla*. PhD thesis, Univ. Paris XI.
- SADEK D., BRETIER P. & PANAGET E. (1997). Artimis : Natural dialogue meets rational agency. In *IJCAI* (2), p. 1030–1035.
- SANSONNET J.-P., LE GUERN K. & MARTIN J.-C. (2005). Une architecture médiateur pour des agents conversationnels animés. In *Actes de WACA'01*, p. 31–39.
- SEARLE J. R. (1969). *Speech Acts : An essay in the Philosophy of language*. Cambridge, new edition.

¹⁰<http://www.limsi.fr/~jps/research/daft/index.html>