

Détection automatique d'actes de dialogues par l'utilisation d'indices multi-niveaux

Sophie Rosset, Delphine Tribout
LIMSI - CNRS
F-91403 Orsay Cedex
{rosset, tribout}@limsi.fr

Mots-clefs : actes de dialogue, dialogue homme homme, détection automatique, indices multiniveaux

Keywords: dialog acts, human human dialog, automatic detection of dialog acts, mulilevel information

Résumé Ces dernières années, il y a eu de nombreux travaux portant sur l'utilisation d'actes de dialogue pour caractériser les dialogues homme-homme ou homme-machine. Cet article fait état de nos travaux sur la détection automatique d'actes de dialogue dans des corpus réels de dialogue homme-homme. Notre travail est fondé essentiellement sur deux hypothèses . (i) la position des mots et la classe sémantique du mot sont plus importants que les mots eux-mêmes pour identifier l'acte de dialogue et (ii) il y a une forte prédictivité dans la succession des actes de dialogues portés sur un même segment dialogique. Une approche de type *Memory Based Learning* a été utilisée pour la détection automatique des actes de dialogue. Le premier modèle n'utilise pas d'autres informations que celles contenus dans le tour de parole. Dans les expériences suivantes, des historiques dialogiques de taille variables sont utilisés. Le taux d'erreur de détection d'actes de dialogue est d'environ 16% avec le premier modèle et descend avec une utilisation plus large de l'historique du dialogue à environ 14%.

Abstract Recently there has been growing interest in using dialog acts to characterize human-human and human-machine dialogs. This paper reports on our experience in the annotation and the automatic detection of dialog acts in human-human spoken dialog corpora. Our work is based on two hypotheses: first, word position is more important than the exact word in identifying the dialog act; and second, there is a strong grammar constraining the sequence of dialog acts. A memory based learning approach has been used to detect dialog acts. In a first set of experiments only the information contained in each turn is used and in a second set, different histories of the dialogue are used. A dialog act error rate of about 16 % is obtained for the simplest model. Using other informations, such as history of the dialog, the results grow up to 14%.

1 Introduction

Afin de saisir la complexité de dialogues homme-homme collectés dans des services d'appels, il semble intéressant d'explorer et de corrélérer différents types d'information, disponibles sous

la forme d'annotations faites à différents niveaux : lexical, sémantique et fonctionnel. D'autre part, il peut être utile de modéliser la structure du dialogue afin d'utiliser cette information dans des systèmes de dialogue. Une analyse souvent effectuée sur les dialogues concerne les actes de dialogue. Les actes de dialogues sont en quelque sorte des unités fonctionnelles abstraites qui décrivent les actions des locuteurs tout en généralisant les variations de forme et de contenu des énoncés. Par exemple, on considère que les assertions (*assert*), les demandes d'informations (*information-request*), les marques d'accord (*acknowledgement*) sont des actes de dialogue qui permettent d'approcher ce que les locuteurs souhaitent accomplir par leur parole. À la base de ce courant, on trouve l'idée selon laquelle *dire c'est faire*¹, c'est-à-dire selon laquelle tout acte d'énonciation serait la réalisation d'un acte social. Cette conception de la parole vient de philosophes du langage ((Austin J. L., 1962),(Searle J. R., 1969)) qui considèrent le dialogue comme un lieu d'interaction sociale et la parole comme un moyen d'(inter)action. Austin considère ainsi qu'une énonciation, outre un contenu explicite, permet également d'accomplir un acte et a donc à ce titre une fonction pragmatique. L'idée d'Austin est en outre que ces fonctions des énoncés peuvent être étudiées indépendamment de leur structure syntaxique mais selon un certain contexte. Plusieurs travaux récents sont fondés sur l'idée que les actes de dialogue sont une bonne façon de caractériser les dialogues, tant dans les interactions homme-homme que homme-machine. Pour exemple, nous pouvons citer les travaux de (Cattoni R. et al., 2001), de (Di Eugenio B. et al., 1998) ou encore ceux de (Isard A., Carletta J. C., 1995). De nombreuses taxonomies d'actes de dialogue ont donc été établies (Traum D., 2000). Pour ce qui concerne les systèmes de dialogue et les annotations de corpus de dialogue homme-homme et homme-machine, une taxonomie fréquemment utilisée et largement répandue est celle de DAMSL². Quant aux approches pour l'annotation automatique en actes dialogiques, il en existe plusieurs qui diffèrent légèrement. Par exemple, ayant observé dans plusieurs corpus que les différents actes de dialogue sont fortement corrélés à des suites de mots précis (appelés *cue-phrases*), (Hirschberg J. et Litman D. J., 1993) se fondent sur ces indices pour les détecter. Le problème de cette approche est toutefois que ces suites de mots sont fortement dépendantes de la tâche et du domaine. Afin de pallier ce problème, (Reithinger N. et Klesen M., 1997) proposent l'utilisation de n-gramme de mots. (Samuel K. et al., 1998), quant à eux, se situent à l'intersection de ces deux approches et utilisent les suites de mots et un sous-ensemble d'indices dialogiques modélisés par des n-gramme de mots. Il est toutefois difficile de ne pas constater que, en règle générale, la relation entre les actes de dialogue et les mots n'est pas univoque. Par exemple, un simple mot comme *oui* peut correspondre à différents actes de dialogue comme une réponse à une question, la confirmation d'une information, un backchannel... D'un autre côté, un acte de dialogue comme une assertion peut correspondre à plusieurs mots ou suites de mots comme *ma date de naissance est le 31/08/70* ou *68 euros 50...* Afin de réduire autant que possible la dépendance à la tâche tout en gérant ces correspondances multiples, nous avons cherché à élaborer une méthode de détection des actes de dialogue sans utilisation explicite du lexique, notre hypothèse étant que cette information n'est pas strictement indispensable.

Dans cet article nous présentons donc notre méthodologie pour la détection automatique des actes de dialogue. Puis nous présentons les différentes expériences que nous avons menées et qui nous ont permis d'améliorer la détection automatique des actes de dialogue grâce à la prise en compte et l'intégration d'indices dialogiques supplémentaires.

¹Ceci fait référence au titre de la version française de (Austin J. L., 1962).

²Cette taxonomie a été utilisée et adaptée dans nombre de projets. Le projet européen et américain AMITIÉS (Automated Multilingual Interaction with Information and Services) par exemple s'est fondée sur elle pour proposer une méthode d'annotation des dialogues sur différents niveaux

nombre de dialogues	134
nombre de tours	4273
nombre moyen de tours/dialogue	32
nombre de segments dialogiques	5623
nombre moyen de segments dialogiques/dialogue	42
nombre moyen de segments dialogiques/tour	1.3
nombre de mots distincts	1976
nombre total de mots	40494

Table 1: Descriptif du corpus GE_fr

2 Corpus et méthodologie

2.1 Corpus utilisé

Dans ce travail, le corpus utilisé (cf. tableau 1) consiste en une série de dialogues homme-homme en français, enregistrés dans un centre d'appel d'un service de prêts bancaires (GE_fr). Ces dialogues couvrent une grande variété de thèmes comme la demande d'informations (limites de crédits possibles, informations sur le disponible), le passage d'ordres (modification des limites de crédits, changement des mensualités...), la gestion de compte (ouverture et fermeture, modification des informations personnelles)... Le corpus au complet est constitué de 134 dialogues. Ces dialogues sont transcrits avec l'outil d'alignement transcription/signal Transcriber. Ce corpus a été divisé en trois parties pour l'apprentissage (94 dialogues, 2923 tours de parole et 3912 segments dialogiques), le développement (22 dialogues, 687 tours de parole et 884 segments dialogiques) et le test (18 dialogues, 663 tours de parole et 827 segments dialogiques). Ce corpus a par ailleurs été tagué en entités spécifiques : entités nommées (personne, lieu, date etc.), entités dépendantes de la tâche (i.e. entités nommées faisant appel à une connaissance spécifique du domaine ; par exemple numéro de compte, adresse, montant disponible sur un compte etc.).

2.2 Principes d'annotation

Ce corpus a été annoté avec le schéma d'annotation dialogique proposé dans le cadre du projet AMITIÉS et fondé sur la taxonomie de DAMSL (Hardy H. et al., 2002). Les annotations devant permettre de décrire et résumer l'intention du locuteur, huit classes ont été établies afin d'obtenir une annotation fine et sur différents niveaux. La taxonomie obtenue est la suivante :

- **Classe 1 *Information Level*** : permet d'annoter un énoncé dans son rapport à la réalisation de la tâche. Les tags possibles sont : *Communication-mgt*, *Out-of-topic*, *Task*, *Task-management-Completion*, *Task-management-Order*, *Task-management-Summary*, *Task-management-System-Capabilities*.
- **Classe 2 *Statement*** : permet d'annoter les énoncés déclaratifs ayant un contenu informatif explicite. Les tags possibles sont : *Assert*, *Commit*, *Explanation*, *Expression*, *ReExplanation*, *Reassert*.
- **Classe 3 *Conventional*** : permet de noter les aspects conventionnels d'un dialogue homme-homme, comme les ouvertures et fermetures. Les tags possibles sont : *Closing*, *Opening*.

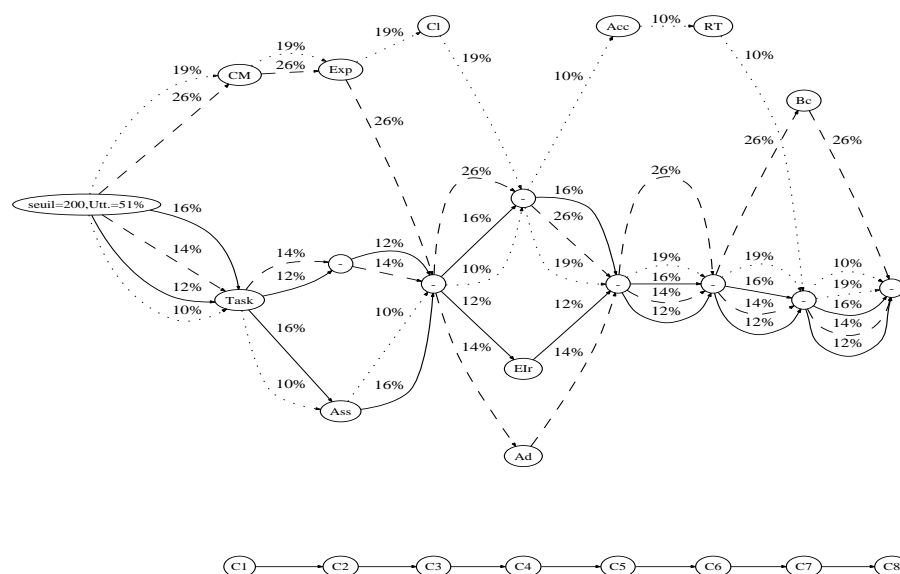


Figure 1: Les successions les plus fréquentes d'ADs

- **Classe 4 Influence on Listener** : permet de rendre compte de l'intention du locuteur en tant qu'agissant sur le déroulement du dialogue. Les tags possibles sont : *Action-directive, Explicit-Confirm-request, Explicit-Info-request, Implicit-Confirm-request, Implicit-Info-request, Offer, Open-Option, Re-Action-directive, Re-Confirm-request, Re-Info-request, Re-Offer*.
- **Classe 5 Agreement** : permet de spécifier l'accord ou le désaccord du locuteur avec ce qui précède. Les tags possibles sont : *Accept, Accept-part, Maybe, Reject, Reject-part*.
- **Classe 6 Answer** : permet de préciser si l'énoncé en question constitue une réponse à un énoncé précédent. Le tag utilisé est alors : *True*.
- **Classe 7 Understanding** : permet de noter les degrés de compréhension du locuteur. Les tags possibles sont : *Backchannel, Completion, Correction, Non-understanding, Repeat-rephrase*.
- **Classe 8 Communicative Status** : permet d'annoter les apartés les interruptions, les changement de sujet... Les tags possibles sont : *AbandStyle, AbandTrans, AbandChangeMind, AbandlossIdeas, Interrupted, Self-talk*.

Les actes de dialogue couvrant différents aspects conversationnels, un même segment dialogique peut contenir plusieurs actes de dialogue. Par conséquent plusieurs de ces classes peuvent être sélectionnées pour décrire un même segment dialogique. Chaque segment dialogique peut en effet être catégorisé selon son niveau informationnel ainsi que selon son aspect conventionnel, son influence sur la suite du dialogue... Ceci implique qu'un segment dialogique peut potentiellement recevoir une étiquette de chacune de ces classes. Par exemple, le segment dialogique *A for Alpha* est annoté avec l'étiquette *Explicit-Confirm-request* de la classe *Influence on Listener* et *Non-understanding* de la classe *Understanding*.

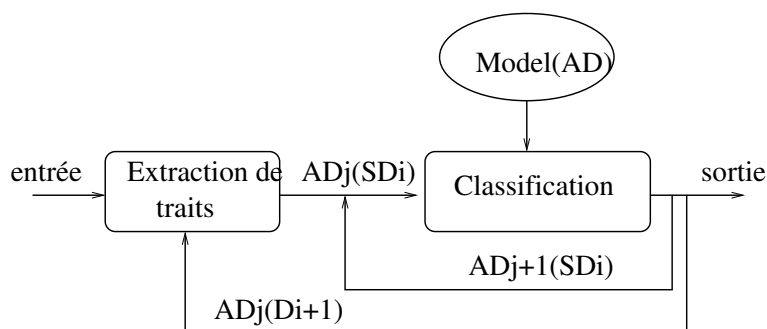


Figure 2: Raisonnement fondé sur la similarité et utilisation des hypothèses précédentes

2.3 Méthodologie pour l'annotation automatique

Pour effectuer l'annotation automatique nous avons représenté chaque énoncé du corpus d'apprentissage et du corpus de développement par un vecteur puis effectué des comparaisons. Pour cela nous avons choisi une approche de type *Memory Based Learning* car elle fonctionne bien sur de petites quantités de données. En outre, différentes études (van den Bosch A. et al., 2001; Daelemans W. et al., 1999) ont montré qu'elle était particulièrement bien adaptée au traitement automatique de la langue. Nous avons utilisé l'implémentation IB1-IG du logiciel Timbl (Daelemans W. et al. 2003) avec une distance de manhatan. Dans cette métrique, la distance entre deux objets est simplement la somme de la différence entre les différents traits de ces objets. Le principe de cette approche est relativement simple : il s'agit de comparer le vecteur entrant à l'ensemble des vecteurs du modèle et d'assigner à celui-ci la classe du vecteur du modèle dont il est le plus proche. Le corpus d'apprentissage a servi de modèle de vecteurs et le corpus de développement a constitué l'ensemble des vecteurs à classer. Les traits choisis pour la construction des vecteurs sont l'identité du locuteur (Client ou Agent), le nombre de segments dialogiques dans le tour considéré, les premiers mots du segment annoté et les tags des huit classes définies. Pour ce qui est de l'utilisation des mots comme traits, notre hypothèse étant que les premiers mots sont plus importants que l'ensemble des mots, seuls les premiers mots de chaque segment dialogique ont été utilisés. Quant aux tags, si aucune étiquette d'une classe considérée n'est pertinente pour le segment alors la classe est représentée par le tag "NA" (not applicable) de manière à avoir toujours le même nombre de traits dans un vecteur. Ainsi, les énoncés du corpus d'apprentissage constituant le modèle de vecteurs ont été représentés de la façon suivante (pour un nombre de mots égal à 4) :

pour un énoncé tel que

Agent: *donnez moi votre numéro de compte*

ayant pour annotation les étiquettes :

DAs: *information-level=Task ; influence-on-listener=Action-directive*

le vecteur correspondant est :

Agent 1 donnez moi votre numéro Task NA NA Action-directive NA NA NA NA

De la même façon, l'énoncé

Client : alors [numerique] [numerique] [numerique] [numerique] [numerique] [numerique]

ayant pour annotation les étiquettes :

DAs: *information-level=Task ; statement=Assert ; agreement=Accept ; answer=true*

est représenté par le vecteur :

Client 1 alors [numerique] [numerique] [numerique] Task Assert NA NA Accept true NA NA

Alors qu'*a priori* la combinatoire des tags est relativement importante, seulement 197 combinaisons différentes sont retrouvées dans le corpus d'entraînement. Il y a donc un facteur de prédictivité relativement important dans la succession des classes et des étiquettes sélectionnées. Ceci est illustré par la figure 1 qui peut être vue comme une grammaire de succession d'étiquettes dialogiques. Sur cette figure, les six successions les plus fréquentes d'étiquettes dialogiques sont représentées. Elles couvrent 51% des séquences du corpus d'entraînement. Ainsi, si Task est sélectionné pour la classe 1 (52%) alors la classe 2 recevra soit NA (26%) soit Assert (26%) et la classe 3 NA. Ceci montre qu'il semble y avoir un facteur de prédictivité relativement important dans la succession des classes et des étiquettes sélectionnées. C'est pourquoi l'annotation des segments dialogiques du corpus à annoter est réalisée en huit étapes : une pour chacune des huit classes d'annotation. A chaque étape un tag est affecté à la classe correspondante et celui-ci est ajouté au vecteur lors de l'étape suivante. En outre, l'annotation dialogique est effectuée en tenant compte de l'ensemble du tour de parole et dans l'ordre des segments composant un tour. C'est-à-dire que pour chaque tour de parole, le premier segment dialogique est d'abord annoté en huit étapes, puis le second segment est ajouté au premier et annoté à son tour en huit étapes. La méthode consiste donc pour le système à extraire du tour de parole donné en entrée les traits retenus (identité du locuteur, nombre de segments dialogiques, N premiers mots) et à les placer dans un vecteur ($SD_1(AD_i)=[\text{SpkrId}, \#\text{Utt.}, w_{-1}, w_{-2}, AD_{i-1}]$). L'assignation d'une étiquette dialogique à ce vecteur (la classification) est faite en comparant ce vecteur à l'ensemble des vecteurs du modèle, et ceci à huit reprises (une pour chacune des classes d'annotation) en ajoutant à chaque étape de la classification l'étiquette déterminée à l'étape précédente. Dès que le premier segment dialogique de l'énoncé en cours de traitement est annoté, si cet énoncé comporte un ou plusieurs autres segments dialogiques, les N premiers mots du segment suivant sont ajoutés au vecteur initial et la classification s'effectue de la même façon. Ainsi le nouveau vecteur est composé de : $SD_i(SD_{i-1} + AD_i + w_{1SD_i}, w_{2SD_i}, \dots)$. Cette méthode utilisée pour annoter automatiquement les énoncés est représentée par la figure 2.

3 Expériences et résultats

3.1 Première expérience : modèle de base

La première expérience, effectuée sur le corpus d'apprentissage, a consisté à faire varier le nombre de mots donnés en entrée : (1) les 4 premiers mots du premier segment dialogique, (2) les 4 premiers mots du premier segment dialogique et les 2 premiers mots du segment pour les segments dialogiques suivants, (3) les 2 premiers mots du premier segment dialogique et les 2 premiers mots du segment pour les segments dialogiques suivants. Les résultats sont présentés dans le tableau 2. Par ailleurs, un examen des poids attribués par TiMBL à chaque trait des vecteurs montre que les mots sont dotés d'un faible poids, ce qui indique qu'ils ne constituent pas le critère le plus pertinent pour catégoriser les vecteurs entrant. Cette observation va bien dans le sens de notre hypothèse lexicale. Toutefois, un examen plus attentif des résultats montrent que certains actes de dialogues sont plus dépendants du lexique que d'autres. Le tableau 3 permet de mettre cette constatation en évidence.

Data	# dial	#seg. dial.	#tour	%erreur	exp.
GE_fr dev	22	884	687	14.0	4words
				13.2	4+2words
				13.0	2+2words
GE_fr Test	18	827	663	16.5	4words
				16.2	4+2words
				17.2	2+2words

Table 2: Taux d'erreur de détection d'actes de dialogues sur corpus developpement et test.

DA	GE_fr dev
Response-To	52% (148)
Backchannel	12.5% (161)
Accept	46.9% (147)
Assert	25.9% (243)
Expression	7.9% (378)
Comm-mgt	8.3% (420)
Task	12.4% (427)

Table 3: Taux d'erreur sur les 7 tags les plus fréquents.

3.2 Réduction de la variation lexicale

Un prétraitement des énoncés a ensuite été effectué pour tenter de réduire la variation de certains énoncés jugés équivalents au niveau sémantique et fonctionnel. L'hypothèse sous-jacente à cette tentative est que l'annotation automatique gênée par certaines variations formelles d'énoncés pourtant équivalents du point de vue du sens pourrait être améliorée si ces variations étaient neutralisées. Ainsi pour la conjugaison des verbes la variation en temps gêne souvent l'annotation. Les formes "je voudrais", "je voulais", "j'aurais voulu", par exemple sont à peu près équivalentes dans le contexte de demande d'information. Nous avons donc décidé de réduire ces variations à une même forme neutralisée "*vouloir", que nous distinguons de l'infinitif avec le signe "*". Les neutralisations concernent également certains morceaux d'énoncés récurrents que nous avons jugés équivalents comme "je vous écoute", "c' est à quel propos", "c' est pourquoi", "c' est à quel sujet"... qui sont tous une façon pour l'agent d'inviter le client à exposer le sujet de son appel, et que nous avons donc neutralisés en *invite. Ces énoncés neutralisés ont été déterminés après l'étude des dialogues du corpus d'apprentissage et concernent notamment les différentes façons de remercier ("je vous remercie", "c' est moi qui vous remercie", "merci beaucoup", "merci bien"...), de demander ("vous pouvez me donner", "vous pouvez me rappeler", "donnez-moi", "rappelez-moi"...) d'ouvrir la conversation ("bonjour", "bonsoir", "allo")... La réduction de ces variations lexicales a été effectuée automatiquement sur l'ensemble du corpus en utilisant l'outil de détection en entités nommées qui a été étendu. Cette étape a permis de faire passer la taille du lexique de 1976 mots à 1649 mots. Les mêmes expériences que précédemment ont ensuite été menées sur ce corpus et les résultats obtenus dans ces conditions montrent que la réduction de la variation lexicale améliore l'annotation automatique. Le tableau 4 donne les résultats obtenus avec ce second modèle.

Exp.	%erreur Dev	%erreur Test.
4words	14.0	16.4
4+2words	12.9	16.2
2+2words	12.7	16.7

Table 4: Taux d’erreur de détection d’actes de dialogues sur corpus développement et test après réduction de la variation

# Seg. Dial	%erreur Dev.	%erreur Test
1 SD	13.2	16.8
2 SD	9.2	16.0
3 SD	22.4	19.2

Table 5: Taux d’erreur de détection d’actes de dialogues sur corpus développement et test par segment dialogique

3.3 Utilisation des historiques

Des expériences supplémentaires ont ensuite été menées en jouant sur l’historique du dialogue afin de voir si cela améliorerait la détection automatique des actes de dialogue. Partant des résultats précédemment obtenus, deux hypothèses ont servi de base à ces expériences. La première est que s’il existe plusieurs segments dialogiques au sein d’un tour, ceux-ci entretiennent des relations entre eux et sont organisés selon certains principes. La seconde est que le dialogue constituant une alternance de tours de parole se répondant les uns aux autres, les actes de dialogue d’un tour sont en partie conditionnés par les actes de dialogue du tour précédent. Les expériences suivantes ont donc été réalisées afin de voir si ces deux hypothèses se vérifiaient.

En ce qui concerne la première hypothèse, nous avons constaté à partir des expériences précédentes qu’au sein d’un tour de parole les résultats de l’annotation automatique étaient systématiquement meilleurs pour le deuxième segment dialogique que pour le premier, et tout aussi systématiquement, nettement moins bons pour le troisième segment³ ainsi que le montre le tableau 5. Le premier segment dialogique, qui est ajouté au vecteur lorsqu’on annote le suivant, semble donc aider à déterminer le second. Celui-ci serait donc étroitement lié au premier et constituerait une sorte de prolongement ou de complément du premier. En revanche, pour le troisième segment les segments précédents qui sont eux aussi ajoutés au vecteur ne semblent pas aider à sa détection. De cette constatation nous avons émis l’hypothèse que si le second segment dialogique d’un tour de parole est dans la continuité du premier, le troisième semble au contraire être en rupture avec eux. Nous avons donc essayé d’annoter automatiquement les segments dialogiques sans tenir compte des deux premiers segments pour annoter le troisième, c’est-à-dire sans mettre les deux premiers segments dans le vecteur du troisième. Avec cette méthode les résultats du troisième segment sont nettement meilleurs comme le montre le tableau 6 (exp. 1). Ceci semble donc confirmer notre hypothèse de départ selon laquelle les segments dialogiques d’un tour de parole entretiennent entre eux des relations qui ne sont pas toutes de la même nature.

En ce qui concerne la seconde hypothèse, les tours de parole se répondant les uns les autres, il nous a semblé intéressant de prendre en compte les informations contenues dans le tour précédent pour déterminer les segments dialogiques d’un tour donné. Toutefois, compte tenu de ce qui vient

³les tours de parole comprenant plus de trois segments dialogiques n’étant pas assez nombreux pour pouvoir émettre une hypothèse à leur sujet l’étude s’est bornée aux trois premiers segments.

Exp.	Dev. %erreur	Test %erreur	Seg. Dial
Exp. 1	13.2	16.8	1
	9.2	16.5	2
	19.7	12.5	3
Exp. 2	10.4	13.7	1
	10.4	15.6	2
	25	14.2	3
Exp. 3	10.4	13.7	1
	10.4	15.6	2
	15.8	13.3	3

Table 6: Taux d'erreur de détection d'actes de dialogues sur corpus developpement et test avec variation sur les historiques

d'être exposé sur les relations que semblent entretenir les segments dialogiques au sein d'un même tour de parole, il nous a semblé que toutes les informations du tour précédent n'étaient sans doute pas pertinentes et que seules les informations relatives au dernier segment dialogique devait avoir une influence sur les segments dialogiques du tour suivant. Nous avons donc ajouté au vecteur les annotations du dernier segment du tour précédent. Les résultats concernant le premier et le deuxième segments ont ainsi été améliorés, mais pas ceux du troisième qui se sont plutôt dégradés (cf. tableau 6 (exp. 2)).

Compte tenu de ces résultats et de ceux obtenus par l'expérience précédente, nous avons essayé de mêler les deux méthodes afin d'améliorer encore l'annotation automatique. Ainsi, pour le premier et le deuxième segments d'un tour de parole les annotations du dernier segment dialogique du tour précédent ont été ajoutées au vecteur, tandis que pour le troisième segment ni les deux segments précédents ni le dernier segment du tour précédent n'ont été pris en compte. Avec cette méthode "mixte", les résultats ont ainsi été améliorés pour tous les segments (cf. tableau 6 (exp. 3)).

4 Conclusion et Perspectives

Ces travaux réalisés dans le cadre de l'annotation dialogique automatique ont permis de mettre en avant le fait que les tours de parole ne sont pas des suites d'énoncés anarchiques mais sont au contraire structurés selon certains principes. La succession des segments dialogiques qui les composent semble en effet organisée : le premier segment dialogique d'un tour de parole semble être lié au dernier segment du tour précédent, le deuxième segment d'un tour semble lié au premier, tandis que le troisième semble indépendant des précédents. En outre, l'étude des dialogues dans leur entier a également fait ressortir une certaine structure du dialogue.

En ce qui concerne l'annotation automatique d'actes de dialogue, les différentes expériences menées ont contribué à son amélioration. Toutefois, celle-ci peut encore être améliorée davantage et d'autres méthodes sont à envisager. Nous envisageons notamment d'utiliser des informations sémantiques, disponibles actuellement sous la forme d'annotations appliquées aux mêmes données, afin de voir si elles ne permettraient pas une meilleure détection des actes de dialogue. Une autre méthode que nous envisageons est d'effectuer une classification des énoncés du corpus d'apprentissage et de faire pour chacune classes dégagées des modèles propres. Ceux-ci seraient

ainsi plus spécifiques et les nouveaux énoncés pourraient alors être annotés en fonction du modèle qui leur correspond le mieux. Enfin, les résultats présentés font tous l'assomption que le nombre de segments dialogiques par tour de parole est connu ainsi que les frontières elles-mêmes. Il est bien entendu que dans le cadre d'un système automatique, ces informations doivent être estimées. Une étude précédente (Rosset S. et Lamel L., 2004) avait montré qu'il était possible de prédire de manière raisonnable le nombre de segment dialogique dans un énoncé. Des taux d'erreurs sur cette tâche de l'ordre de 12% ont été rapportés. Environ 5 à 7% des tour de parole présentaient une insertion ou une suppression de frontière. Ces modèles pour la détection de frontières de segments dialogiques doivent donc également être intégrés au système de détection d'actes de dialogue pour avoir un système entièrement automatique.

Références

- J. L. Austin (1962), *How to do thing with words*, Oxford: Clarendon Press.
- R. Cattoni, M. Danieli, A. Panizza, V. Sandrini, C. Soria (2001), Building a corpus of annotated dialogues: the ADAM experience, *Corpus Linguistics* 2001.
- W. Daelemans, J. Zavrel, K. van der Sloot, A. van den Bosch (2003), TiMBL: Tilburg Memory Based Learner, v5.0, Reference Guide, *ILK Technical Report ILK-03-10*, (<http://ilk.kub.nl/software.html#timbl>)
- W. Daelemans, A. van den Bosch, J. Zavrel (1999), Forgetting exceptions is harmful in language learning, *Machine Learning* Vol 34:11-43.
- B. Di Eugenio, P. W. Jordan, J. D. Moore, R. H. Thomason (1998) An empirical investigation of collaborative dialogues, *Actes de ACL, COLING*.
- H. Hardy, K. Baker, H. Bonneau-Maynard, L. Devillers, S. Rosset, T. Strzalkowski (2002), Semantic and Dialogic annotation for Automated Multilingual Customer Service, *Actes de ISLE workshop*.
- J. Hirschberg, D.J. Litman (1993), Empirical Studies on the Disambiguation of Cue Phrases, *Computational Linguistics* Vol. 19(3):501-530.
- A. Isard, J.C. Carletta (1995), Replicability of transaction and action coding in the map task corpus, *Actes de AAAI Spring Symposium: Empirical Methods in Discourse Interpretation and Generation*.
- N. Reithinger, M. Klesen. (1997), Dialogue act classification using language models, *Actes de Eurospeech'97*
- S. Rosset et L. Lamel (2004), Automatic Detection of Dialog Acts Based on Multi-level Information, *Actes de ICSLP'04*
- K. Samuel, S. Carberry, K. Vijay-Shanker (1998), Dialogue act tagging with transformation-based learning, *Actes de COLING-ACL*
- J. R. Searle (1969), *Speech acts*, Cambridge University Press.
- D. Traum (2000), 20 Questions on Dialog Act taxonomies, *Journal of Semantics*, Vol. 17(1):7-30.
- A. van den Bosch, E. Krahmer, M. Swerts (2001), Detecting problematic turns in human-machine interactions: Rule-induction versus memory-based learning approaches, *Actes de ACL'01*