

L'analyse sémantique de dialogues oraux transcrits

Mohamed Gouiaa¹, Sylvain Delisle² Bernard Moulin^{1,3}

(1) Département d'informatique, Université Laval, Québec, Canada

(3) Centre de recherche en géomatique de l'Université Laval

moulin@ift.ulaval.ca

(2) Département de mathématiques et d'informatique

Université du Québec à Trois-Rivières, Québec, Canada

Sylvain_Delisle@uqtr.quebec.ca

Résumé

Nous présentons les principaux résultats obtenus à ce jour dans le cadre du projet MAREDI qui vise à développer un système de traitement de la langue naturelle permettant d'analyser des transcriptions de dialogues oraux et de générer un modèle conceptuel de la conversation. Nous discutons principalement des aspects touchant l'analyse sémantique, en l'occurrence le rôle qu'y jouent les actes de discours et l'analyse casuelle, tout en présentant brièvement l'architecture globale du système et les caractéristiques de ses différentes composantes.

1. Introduction

Les discours résultent de productions linguistiques créées de façon coopérative par des agents locuteurs ou logiciels. Les agents choisissent la forme et le contenu de leurs interventions en fonction du contexte conversationnel, de leurs objectifs et des relations sociales qu'ils entretiennent avec les autres agents participant à la conversation. Dans une conversation, un agent interagit avec les autres locuteurs en accomplissant différents types d'actes de communication, tels que des énonciations (appelés actes de discours) et des gestes. Dans le projet MAREDI (MARqueurs et REprésentation des DIscours), nous visons à développer une approche et un outil d'analyse de discours qui s'appuie sur la détection et l'interprétation de marques de surface afin d'élaborer un modèle conceptuel de discours sous la forme d'un ensemble d'états mentaux structurés. Pour la représentation du modèle conceptuel du discours, nous adoptons le modèle de Rousseau *et al.* (1994) qui propose une approche multi-agent pour modéliser et simuler une conversation entre des agents logiciels capables de communiquer entre eux, de raisonner sur leurs états mentaux, de planifier et d'exécuter des actions non-linguistiques et des actes de discours. Le contenu conceptuel des actes de discours accomplis par les agents locuteurs sont modélisés par des objets conversationnels qui correspondent à certains états mentaux que les agents désirent échanger. Ces objets conversationnels, et d'autres structures de connaissances spéciales touchant le maintien de la conversation et la gestion de l'initiative, sont agrégées dans une mémoire persistante de la conversation gérée par un agent particulier : l'agent de conversation. Les agents locuteurs peuvent raisonner sur le contenu de leur propre modèle mental et du modèle de la conversation.

Nous avons travaillé sur un corpus de dialogues de conception de figures géométriques, constitué par Ozkan (1994) et analysé par Colineau (1994). Ce corpus met en scène deux locuteurs, un instructeur et un manipulateur placés en situation de conception de figures dans deux salles séparées. Les locuteurs communiquaient entre eux grâce à des micros et partageaient les mêmes informations sur leurs écrans (les actions exécutées par l'un étaient visibles par l'autre). L'instructeur (Inst) avait sous les yeux la scène à représenter et donnait des directives au manipulateur (Man) qui ignorait ce qu'il devait dessiner. La Figure 1 présente un

dialogue, extrait d'un corpus, dans lequel nous remarquons un certain nombre de distorsions : répétitions de mots, interruption des énoncés, énoncés incomplets, énoncés bruités, etc. En accord avec le modèle de Rousseau *et al.* (1994), les interactions des locuteurs sont modélisées en tant qu'échanges d'objets conversationnels (OC). Un objet conversationnel est un état mental (croyance, but, souhait, etc.) transféré d'un locuteur à un ou plusieurs autres locuteurs au cours de l'accomplissement d'un acte de discours. Le locuteur se positionne relativement à un état mental en exécutant des actions comme "proposer", "accepter" et "refuser": on parle ainsi de positionnement du locuteur par rapport à cet état mental.

Inst (i1)	euh main'nant tu vas prendre un petit un p'tit carré
Inst (i2)	et le mettre en bas à gauche
Inst (i3)	pour faire une maison pour euh
Man (m4)	geste (exécution)
Inst (i5)	maintenant tu en prends un autre
Man (m6)	geste (prise et déplacement)
Man (m7)	avec un toit aussi ?
Inst (i8)	non non

Figure 1. Extrait d'un corpus : dialogue orienté tâche (Inst : instructeur ; Man : manipulateur)

Notre approche est fondée sur l'hypothèse que diverses marques trouvées dans un discours peuvent être utilisées pour orienter l'analyse et identifier plus spécifiquement les catégories des actes de dialogue (Colineau 1997). Diverses équipes de recherche ont adopté cette hypothèse comme en témoignent par exemple les recherches récentes sur le "tagging" d'actes de dialogue (Samuel *et al.* 1998), l'étude des actes de dialogues qui servent à structurer les discours (Jurafsky *et al.* 98) et l'élaboration de lexiques de marqueurs de discours (Stede *et al.* 1998). Les principales étapes de l'analyse sont : 1) découper le texte initial en énoncés; 2) supprimer les bruits tels que "Euh"; 3) identifier les marques dans les énoncés (telles que "maintenant", "aussi") et les extraire du texte; 4) procéder à l'analyse syntaxique des énoncés, qu'ils soient bien formés, mal formés ou fragmentaires; 5) pour chaque énoncé, identifier la catégorie de l'acte de dialogue (c'est une analyse pragmatique qui tient compte de nombreux actes de discours indirects trouvés dans les dialogues oraux); 6) en se basant sur la catégorie du discours, identifier le positionnement du locuteur et le type d'état mental (pour ce faire, une table de correspondance est utilisée entre le type d'acte de dialogue et le couple [positionnement, type d'état mental]); 7) procéder à une analyse sémantique du contenu propositionnel du discours en se basant sur les résultats syntaxiques; et 8) créer le modèle conceptuel de la conversation.

La Figure 2 montre l'architecture du système MAREDI. Les transcriptions écrites des dialogues oraux sont initialement traitées par un filtre pré-syntaxique permettant de recueillir un certain nombre de marques (locutions, expressions figées, marques lexicales, etc.). Le texte filtré est ensuite passé à un analyseur syntaxique robuste (Boufaden *et al.* 1998) qui en produit un arbre syntaxique et l'envoie à un filtre post-syntaxique. Ce dernier identifie une autre partie des marques syntaxiques qui sont en rapport avec le verbe : le type du verbe (verbe d'action, verbe d'état), la structure du verbe (le verbe est seul ou conjugué avec un verbe modal), si le verbe est membre de la proposition principale ou bien d'une proposition subordonnée, ainsi que le temps et le mode du verbe. Les deux filtres constituent deux programmes d'appariement qui identifient les marques pertinentes figurant dans l'énoncé ou dans l'arbre syntaxique et ce, en utilisant une table contenant les divers types de marques rencontrées dans nos corpus. À partir des marques pré-syntaxiques (p.ex. particule interrogative) et post-syntaxiques (p.ex forme et modalité verbale), un analyseur neuronal identifie la catégorie de l'acte de dialogue et l'envoie au module intégrateur. L'intégrateur reçoit également l'interprétation sémantique du contenu propositionnel de l'acte produite par l'analyseur sémantique à partir

du résultat d'analyse syntaxique. Le module intégrateur produit une représentation conceptuelle du discours qui sera intégrée dans le modèle de la conversation.

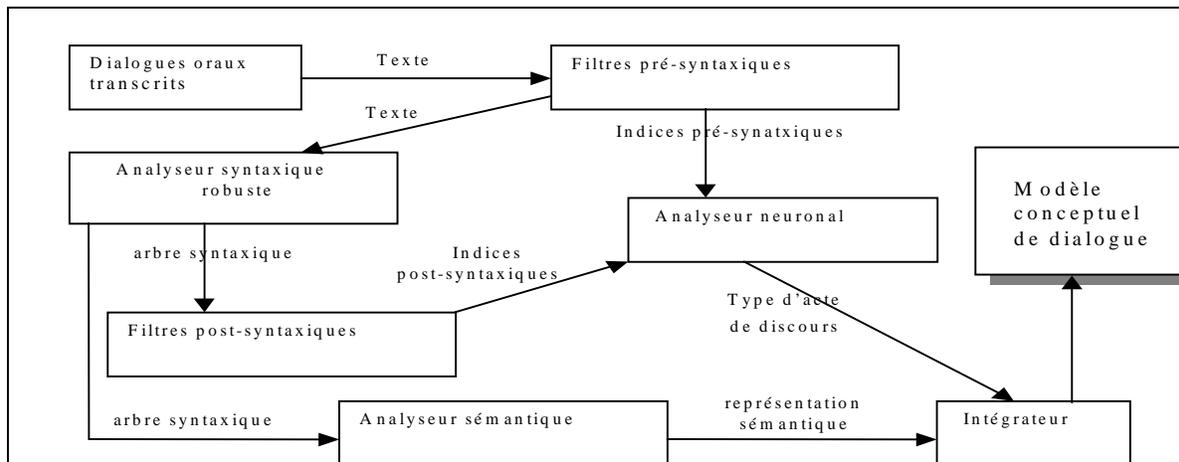


Figure 2. Architecture du système MAREDI

Dans ce qui suit, nous présentons l'analyseur sémantique de MAREDI. Précisons que certains des aspects décrits pourraient être traités davantage du point de vue de la pragmatique. Mais comme notre analyse est faite dans une perspective logico-conceptuelle, nous avons préféré nous situer dans le giron de la sémantique. Les aspects linguistiques (modèles et analyseur) ont été présentés dans Boufaden *et al.* (1998).

2. Vers un modèle conceptuel de la conversation

Le modèle conceptuel de la conversation est obtenu grâce à la contribution de l'analyseur neuronal—détails dans Colineau, 1997—qui détermine le type de l'acte de discours correspondant à l'énoncé analysé, et de l'analyseur sémantique qui fournit la structure sémantique du contenu propositionnel de l'acte de discours.

2.1. Un analyseur sémantique casuel

L'analyseur sémantique casuel (Gouiaa 1998) effectue une analyse casuelle (Delisle *et al.* 1996, Somers 1987) ayant recours aux primitives des approches *templates* et *inference-driven mapping* (Palmer 1990). Deux raisons motivent ce choix—pour un autre point de vue, voir Antoine (1996). La première est que ces approches procèdent d'abord à une analyse syntaxique suivie d'une analyse sémantique dans deux modules séparés, ce qui constitue un avantage du point de vue de la réalisation informatique. La seconde raison est la mise en correspondance directe existant entre les constituants syntaxiques et les cas sémantiques. Nous avons opté pour une représentation de la phrase en fonction d'une composante modale et d'une proposition. La composante modale comporte des informations sur la négation, le temps, le mode (p.ex. indicatif, impératif, conditionnel, subjonctif, infinitif) et la modalité (p.ex. obligation, désir, possibilité, interrogation, etc.). La proposition permet l'identification du verbe (ou, prédicat) et la reconnaissance des relations sémantiques qui lient les syntagmes nominaux et propositionnels (ou, arguments) au verbe. À cela, notre modèle de représentation rajoute les informations concernant les modificateurs qui sont principalement des adjectifs. Nous aboutissons à la fin de l'analyse à une représentation sémantique, basée sur la théorie des cas, qui montre le verbe avec ses cas sémantiques instanciés par les constituants syntaxiques correspondants.

Dans un lexique, chaque verbe se voit attacher *a priori* ses cas sémantiques possibles, ce qui permet de mettre en évidence les relations de sens qui existent entre un prédicat et ses

arguments. Ceci nous a amené à définir un ensemble de cas sémantiques (p.ex. AGENT, OBJET, LOCALISATION, BUT), basé sur Barker *et al.* (1997), ainsi qu'un ensemble de *templates* qui caractérisent la structure casuelle des verbes rencontrés dans les corpus. Le caractère obligatoire ou optionnel des cas est indiqué afin d'apporter un certain degré de flexibilité au niveau des *templates*. Ceci nous permet de traiter des énoncés qui sont similaires (p.ex. "tu prends un carré" et "tu prends un carré du bas de l'écran") et de réduire ainsi le nombre d'entrées, donc de *templates*. Nous définissons également un ensemble de *patrons-types* : des patrons-types de proposition et des patrons-types de modalité. Les patrons-types de modalité permettent de déterminer les éléments constituant la modalité de la phrase. Les patrons-types de proposition permettent de retrouver les éléments syntaxiques ("fillers") qui remplissent les cas sémantiques attachés au verbe. L'unification de l'énoncé, ou plus exactement de son arbre d'analyse syntaxique, avec le *template* qui lui correspond permet d'associer à chaque constituant syntaxique un rôle sémantique. La validité des rôles sémantiques remplis est vérifiée à l'aide d'un certain nombre de contraintes qu'on appelle "restrictions de sélection" comme, par exemple, vérifier si le cas AGENT, qui occupe la position sujet, est un être animé dans "Jean mange une pomme". Pour appliquer ces contraintes, nous décrivons une *grille de types* de concepts (Sowa 1984), qui est une structure de treillis caractérisant notre univers du discours. Cette grille de types de concepts englobe les différents éléments qui figurent dans les corpus de MAREDI, auxquels sont associés des types et des sous-types.

2.2. Évaluation expérimentale de l'analyseur sémantique casuel

L'évaluation porte fondamentalement sur : 1) le pourcentage des différentes structures syntaxiques produites par l'analyseur syntaxique qui sont reconnues par l'analyseur sémantique ; 2) le pourcentage des analyses sémantiques correctes pour les structures syntaxiques reconnues ; 3) le pourcentage des énoncés qui se voient attribuer une analyse sémantique correcte (en accord avec notre modèle théorique). Afin d'effectuer l'évaluation, nous avons analysé deux corpus C11-Homme-chaise et C5-Lampe (Colineau 1994). Le corpus C11-Homme-chaise comporte 106 énoncés et le corpus C5-Lampe comporte 84 énoncés. Les Tableaux 1 et 2 présentent pour chaque type d'énoncés le nombre et le pourcentage correspondant.

Type d'énoncé	Corrects	Const. figées ou gestes	Elliptiques	Interrompus
Nombre	36	54	9	7
Pourcentage	33,92%	50,94%	8,94%	6,60%

Tableau 1. Nombre et pourcentage d'énoncés dans C11-Homme-chaise

Type d'énoncé	Corrects	Const. figées ou gestes	Elliptiques	Interrompus
Nombre	32	40	7	5
Pourcentage	38,09%	47,61%	8,33%	5,95%

Tableau 2. Nombre et pourcentage d'énoncés dans C5-Lampe

À partir de ces deux tableaux, nous distinguons quatre types d'énoncés. Les énoncés corrects contiennent un noyau verbal : ce sont les énoncés autres que ceux qui sont interrompus, elliptiques ou formés uniquement par des constructions figées. Le second type concerne les énoncés composés uniquement de constructions figées ("Const. figées") ou bien d'énoncés gestuels, puis les énoncés elliptiques et finalement les énoncés interrompus. Les énoncés composés par des constructions figées sont les plus nombreux (50,94% dans C11-Homme-chaise et 47,61% dans C5-Lampe). Ensuite viennent les énoncés corrects. Les énoncés elliptiques et interrompus sont les moins nombreux (à peu près 8% et 6% dans les deux corpus). L'analyseur syntaxique génère des structures syntaxiques partielles pour les énoncés elliptiques et les énoncés interrompus. À partir de ces structures partielles, l'analyseur sémantique produit, dans la mesure du possible, des structures d'interprétation sémantique qui sont accompagnées par des messages indiquant les parties manquantes de l'énoncé (en spécifiant les

rôles définis comme obligatoires dans les structures casuelles, et qui n'ont pas été remplis par des *fillers*). Pour les constructions figées, l'analyseur syntaxique recueille les différentes marques présentes dans l'énoncé, identifie (à partir de ces marques) le type d'acte de discours, et génère la structure d'interprétation sémantique correspondante. Dans les Tableau 3 et 4, nous considérons uniquement les énoncés syntaxiquement corrects. Nous remarquons une différence assez importante dans les pourcentages calculés qui est due aux constructions interrogatives qui sont plus nombreuses dans le corpus C5-Lampe, et qui, dans la version actuelle de l'analyseur syntaxique, ne sont pas traités correctement.

Nombre d'énoncés corrects	Énoncés syntaxiques reconnus	% analyseur sémantique
36	33	91,66%

Tableau 3. Pourcentage des énoncés syntaxiques reconnus dans C11-Homme-chaise

Nombre d'énoncés corrects	Énoncés syntaxiques reconnus	% analyseur sémantique
32	22	68,75%

Tableau 4. Pourcentage des énoncés syntaxiques reconnus dans C5-Lampe

Les Tableaux 5 et 6 montrent les pourcentages des analyses sémantiques correctes pour les énoncés syntaxiquement corrects et ce, en comptabilisant le nombre d'énoncés qui ont été traités correctement par l'analyseur sémantique.

Structures syntaxiques reconnues	Analyses sémantiques correctes	% analyseur sémantique
33	27	81,81%

Tableau 5. Pourcentage des analyses sémantiques correctes dans C11-Homme-chaise

Structures syntaxiques reconnues	Analyses sémantiques correctes	% analyseur sémantique
22	17	77,27%

Tableau 6. Pourcentage des analyses sémantiques correctes dans C5-Lampe

Les énoncés dont l'analyse sémantique a échoué sont généralement ceux qui emploient le verbe d'état *être* (p.ex. "*c'est les pieds de la chaise ça*", "*ça sera le dossier*"). Cela est dû à la difficulté que posent les verbes d'état dans une analyse casuelle. Pour le reste des types d'énoncés, les pourcentages de succès de l'analyse sémantique sont présentés dans les deux tableaux suivants :

Types d'énoncés	Constr. figées ou gestes	Elliptiques	Interrompus
Nombre d'énoncés	54	9	7
Structure sémantique	54	0	3
Pourcentage	100%	0%	42,85%

Tableau 7. Pourcentage des analyses sémantiques dans C11-Homme-chaise

Types d'énoncés	Constr. figées ou gestes	Elliptiques	Interrompus
Nombre d'énoncés	40	7	5
Structure sémantique	40	0	2
Pourcentage	100%	0%	40%

Tableau 8. Pourcentage des analyses sémantiques dans C5-Lampe

Les résultats pour les énoncés composés par des constructions figées indique un pourcentage de 100% qui est normal puisqu'aucune structure syntaxique n'est générée pour ce type d'analyse. Le traitement sémantique s'effectue plutôt à partir d'une liste de marques identifiées par l'analyseur syntaxique. L'analyseur sémantique détermine à partir des marques relevées dans l'énoncé le type d'acte de discours approprié, et génère ainsi la structure d'interprétation correspondante. En ce qui concerne les énoncés elliptiques, nous remarquons dans les deux tableaux que le taux de réussite est nul. Cela est dû au fait que ce type de construction n'est pas encore traité par notre analyseur. Des analyses syntaxiques partielles sont

produites pour les énoncés interrompus. À partir de ces analyses, nous avons procédé à l'interprétation sémantique de ces énoncés. Les structures générées sont des structures sémantiques partielles accompagnées par des messages indiquant les rôles sémantiques manquants (grâce aux structures casuelles des verbes employés) avec le type exigé pour l'élément syntaxique (*filler*) de ce rôle.

3. Conclusion et travaux futurs

Dans le contexte du traitement d'énoncés provenant de l'oral, rares sont les analyseurs sémantiques qui abordent l'ensemble des questions considérées ici : cela constitue un des points saillants de notre travail. Comme nous l'avons expliqué, le traitement sémantique ne se limite pas aux énoncés bien formés. Nous avons présenté les deux composantes (analyseur neuronal et sémantique) qui sont chargées d'élaborer le modèle conceptuel de la conversation dans le système MAREDI, en mettant une emphase particulière sur l'analyseur sémantique casuel. Un aspect original de ce système résulte de la complémentarité des trois analyseurs, syntaxique, neuronal et sémantique, pour effectuer une analyse de transcriptions de l'oral et pour élaborer le modèle conceptuel de la conversation. Cette complémentarité est appréciable pour le traitement des énoncés mal formés qui sont si nombreux dans les interactions orales. Le projet MAREDI est en à sa dernière étape : les analyseurs doivent être encore raffinés et intégrés dans un système unique et une évaluation globale de la performance du système doit être faite.

Références

- ANTOINE J-Y. (1996), "Parsing Spoken Language Without Syntax", Proceedings of the 16th International Conference on Computational Linguistics (COLING-96), Copenhagen, Denmark, 5-9 August 1996, 47-52.
- BARKER K., COPECK T., DELISLE S., SZPAKOWICZ S. (1997), "Systematic Construction of a Versatile Case System", Journal of Natural Language Engineering 3 (4), 279-315.
- BOUFADEN N., DELISLE S., MOULIN B. (1998), "Analyse syntaxique robuste de dialogues transcrits: peut-on vraiment traiter l'oral à partir de l'écrit?", Actes de la Cinquième Conférence Annuelle sur le Traitement Automatique des Langues Naturelles (TALN'98), Paris (France), 10-12 juin 1998, 112-121.
- COLINEAU N. (1994), "Vers une compréhension des actes de discours", rapport de DEA en Sciences Cognitives, Institut National Polytechnique de Grenoble (France).
- COLINEAU N. (1997), "Etude des marqueurs discursifs dans un dialogue finalisé", thèse de Doctorat es Sciences Cognitives, Université Joseph Fourier, Grenoble (France).
- DELISLE S., BARKER K., COPECK T., SZPAKOWICZ S. (1996), "Interactive Semantic Analysis of Technical Texts", Computational Intelligence, 12(2), 273-306.
- GOUIAA M. (1998), "Modélisation conceptuelle des conversations à partir de l'analyse de textes de dialogues transcrits: analyseur sémantique et générateur de la structure conversationnelle", mémoire de maîtrise en informatique, Département d'informatique, Université Laval.
- JURAFSKY D., SHRIBERG E., FOX B., CURL T. (1998), "Lexical, Prosodic and Syntactic Cues for Dialog Acts", Proceedings of the Workshop on Discourse Relations and Discourse Markers (COLING-ACL'98), Montréal (Québec, Canada), 15 août 1998, 114-120.
- OZKAN N. (1994), "Vers un modèle dynamique du dialogue : analyse de dialogues finalisés dans une perspective communicationnelle", thèse de Doctorat en Sciences Cognitives, Institut National Polytechnique de Grenoble (France).
- PALMER, M. (1990), Semantic processing for finite domains. Cambridge University Press.
- ROUSSEAU D., MOULIN B., LAPALME G. (1994), "Une approche multi-agents pour modéliser les conversations", Actes des Deuxièmes journées francophones Intelligence Artificielle Distribuée et Systèmes Multi-Agents, Voiron (France), 3-14.
- SAMUEL K., CARBERRY S., VIJAY-SHANKER K. (1998), "Dialogue Act Tagging with Transformation-based Learning", Proceedings of the Conference COLING-ACL'98, Montréal (Québec, Canada), 10-14 août 1998, 1150-1156.
- SOMERS H.L. (1987) Valency and Case in Computational Linguistics. Edinburgh University Press.
- SOWA J. (1984), Conceptual Structures : Information Processing in Mind and Machine, Addison Wesley publishing company, Reading, MA.
- STEDE M., UMBACH C. (1998), "DiMLex : A Lexicon of Discourse Markers for Text Generation and Understanding", Proceedings of the Conference COLING-ACL'98, Montréal (Québec, Canada), 10-14 août 1998, 1238-1242.