



**HAL**  
open science

# Analyse des sentiments dans les conversations humain-agent Vers un modèle des goûts de l'utilisateur

Caroline Langlet

► **To cite this version:**

Caroline Langlet. Analyse des sentiments dans les conversations humain-agent Vers un modèle des goûts de l'utilisateur. Traitement du texte et du document. Télécom ParisTech, 2018. Français. NNT: . tel-02002580v1

**HAL Id: tel-02002580**

**<https://theses.hal.science/tel-02002580v1>**

Submitted on 31 Jan 2019 (v1), last revised 11 Feb 2019 (v2)

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



EDITE - ED 130

## Doctorat ParisTech

### THÈSE

pour obtenir le grade de docteur délivré par

**TELECOM ParisTech**

**Spécialité « Informatique »**

*présentée et soutenue publiquement par*

**Caroline LANGLET**

le 26 septembre 2018

**Analyse des sentiments dans les conversations humain-agent**

**Vers un modèle des goûts de l'utilisateur**

Directeur de thèse : **Catherine Pelachaud**

Co-encadrement de la thèse : **Chloé Clavel**

#### Jury

**Mme Catherine PELACHAUD**, Titre, ISIR, CNRS

**Mme Chloé CLAVEL**, Maître de conférences HDR, LTCI, Paris-Saclay

**M. Thierry POIBEAU**, Professeur, Lattice, CNRS

**M. Bjorn SCHULLER**, Professeur, GLAM, Imperial College of London

**Mme Pascale SÉBILLOT**, Professeur, IRISA, INSA Rennes

**M. Nicolas MAUDET**, Professeur, LIP6, Sorbonne Université

**M. Dirk HEYLEN**, Professeur, HMI, University of Twente

**Mme Marie-Jeanne LÉSOT**, Maître de conférences HDR, LIP6, Sorbonne Université

Directeur de thèse

Co-encadrant de thèse

Rapporteur

Rapporteur

Examineur

Examineur

Examineur

Examineur

**TELECOM ParisTech**

école de l'Institut Mines-Télécom - membre de ParisTech



*À Nicolas,*

*Nani gigantum humeris insidentes*

# Résumé

Cette thèse se situe à la croisée de deux domaines de recherche : celui du *sentiment analysis* et celui des agents conversationnels animés. Les agents conversationnels animés peuvent être définis comme des personnages virtuels ayant la capacité de converser avec un utilisateur humain. Afin d'accroître les compétences communicationnelles de l'agent, il est important que celui-ci soit doté d'une forme d'intelligence socio-émotionnelle. L'agent doit être ainsi en capacité de gérer des signaux socio-émotionnels, tant du côté de la génération que de celui de la détection. Du côté de la génération, de nombreux travaux ont produit des modèles optimisant la production de gestes ou d'expressions faciales pour exprimer soit des émotions soit des attitudes sociales. Du côté de la détection, une majorité des travaux se concentrent sur l'analyse d'indices socio-affectifs non-verbaux (expressions faciales, indices acoustiques). Le contenu verbal et les expressions de sentiment qu'il véhicule restent quant à lui encore partiellement exploité. En effet, les rares études, intégrant un module de détection des sentiments de l'utilisateur dans le cadre de conversations humain-agent, ne prennent pas en compte les spécificités de ce contexte d'interaction. Pour combler cette lacune, notre travail s'intéresse à l'analyse du contenu verbal produit par l'utilisateur et à la manière dont celui-ci réfère ou exprime des sentiments, des affects ou des attitudes. Nous en proposons un modèle de détection au cours d'une interaction multi-modale et en face à face avec un agent conversationnel animé. Pour construire ce modèle, deux questions se sont posées à nous. Dans un premier temps, il nous a fallu identifier, au sein de la vaste classe des expressions de sentiment, celles qui apparaissent comme les plus pertinentes pour l'élaboration des stratégies de communication de l'agent. Dans un second temps, nous avons dû choisir une méthode devant être non seulement opérante pour une analyse à grain fin de ces expressions, mais également adaptable au contexte conversationnel.

Nos contributions s'articulent autour de trois axes. Tout d'abord, nous fournissons une analyse approfondie des expressions de sentiment. Trois unités conversationnelles sont considérées : le tour de parole, la paire adjacente et la séquence thématique. Cette analyse met en évidence un certain nombre de caractéristiques nécessaires au développement d'un ensemble de règles de détection. Ensuite, nous proposons un modèle de détection symbolique intégrant des règles sémantiques et des grammaires formelles. Ce modèle repose sur une analyse ascendante des énoncés – du niveau lexical au niveau phrastique – et se concentre successivement sur trois cadres d'analyse : le tour de parole, la paire adjacente et la séquence thématique. Enfin, nous proposons un protocole d'évaluation pour la validation des règles. Grâce à la création de deux plateformes d'annotation, nous avons pu créer deux jeux d'annotations sur deux corpus différents : un corpus de *small-talk* et un corpus de négociation. Les performances du système ont ainsi pu être évaluées par rapport aux références obtenues.

**Mots-clefs :** Analyse de sentiments, Interactions humain-agent, Agents conversationnels animés.



# Abstract

## **Sentiment Analysis in Human-Agent Conversations : Modelling User's Like and Dislikes**

This thesis is related to two research areas : the sentiment analysis and the embodied conversational agents. Embodied conversational agents can be defined as virtual characters with the ability to converse with a human user. In order to increase the communicative skills, the agent needs a social-emotional intelligence. The agent must be able to manage socio-emotional signals, both for generation and detection. Many works have produced models optimizing the production of gestures or facial expressions expressing either emotions or social attitudes. On the detection side, a majority of the work focuses on the analysis of nonverbal and social-affective features (facial expressions, acoustic features). The verbal content is more and more integrated but remains partially exploited. The rare studies integrating a detection module of user's sentiments in the context of conversations human-agent do not integrate the specificities of this interaction context.

This thesis focuses on the user's verbal expressions of sentiment. It is dedicated to the development of a detection model for face-to-face human-agent conversations. First, our goal is to define, in the broad class of sentiment expressions, those appearing as the most adapted to the agent's communicative goals. Second, we aim to select a method that offers the possibility of a fine-grained analysis and can be suited to the conversational context.

We choose to refer to the model of Martin and White (2005) and to use a method integrating linguistic rules. Our contributions comprise three parts. First, we provide an in-depth analysis of sentiment expressions that focus on three conversational units : the speech turn, the adjacency pair and the topic sequence. This analysis highlights a number of features used for the development of linguistic rules. Then, we propose a system integrating rules and grammars. The system is grounded on an bottom-up analysis starting from the lexical level to reach the phrastic level. Three level are successively analysed : the speech turn, the adjacency pair and the topic sequence. Finally, we propose an evaluation of the rules. We created two annotation platforms for building two annotations sets on two corpora : a *small-talk* corpus and a negotiation corpus. The annotation sets were used for evaluating the performance of the system.

**Keywords :** Sentiment Analysis, Human-Agent Interaction, Embodied Conversational Agents.



# Remerciements

Tout d’abord, je souhaiterais exprimer ma profonde gratitude envers mes deux directrices de thèse : Chloé Clavel et Catherine Pelachaud. Je souhaiterais remercier Chloé pour la confiance et la liberté d’expérimentation qu’elle m’a accordées. Merci d’avoir toujours accueilli mes idées avec un regard critique et une force de proposition permettant de les perfectionner et d’en tirer le meilleur. Merci pour les relectures tard le soir avant les *deadlines* et pour les nombreux retours sur mon travail. Je souhaiterais remercier Catherine pour avoir accordé de l’attention à mon travail malgré l’écart entre nos deux domaines de recherche.

Au terme de ma deuxième année de thèse, j’ai dû faire face à un événement tragique m’obligeant à quitter la *Greta team* pendant presque une année. Je ne peux remercier Catherine et Chloé sans évoquer le soutien, tant moral que matériel, dont elles toutes deux ont fait preuve pendant cette période difficile. Vos encouragements et votre aide m’ont donné la force de poursuivre et d’achever mon travail de thèse. A ce titre, je souhaiterais également remercier Bertrand Grando et Alain Sibille, directeurs de l’école doctorale, pour m’avoir aidé sur le plan administratif à reprendre mes travaux de thèse. J’adresse également des remerciements sincères aux personnels administratifs de IDS (Marie-Laure Chauveaux, Laurence Zelmar et Salimatou Yansane) ainsi qu’à Florence Besnard, pour les nombreux services qu’ils ont pu me rendre pendant ma thèse.

Au cours de mon doctorat, j’ai eu la chance de pouvoir évoluer dans un cadre exceptionnel : la *Greta team*. Une équipe où la joie et le rire (mais aussi les changements de fond d’écran) s’associent au sérieux et à l’esprit critique pour donner une ambiance de travail des plus remarquables. Sont particulièrement ancrés dans ma mémoire les membres de l’équipe telle qu’elle était constituée à mon arrivée : Ken Prepin, Andre-Marie Pez, Pierre Philippe, Yu Ding, Jing Huang, Magalie Ochs. Comment ne pas se souvenir des rires échangés entre Florian Pecune, Brian Ravenet et Mathieu Cholet, des débats animés (et parfois houleux) avec Brice Donval et Kevin Sanlaville, des blagues de linguistes que je partageais avec Sabrina Campano, des discussions toujours passionnantes que j’ai pu échanger avec Nadine Glas (je repense à cette nuit à Xi’An où nous avons parlé pendant des heures) ? *The last but not the least*, je souhaiterais adresser un remerciement particulier à Nesrine Fourati pour sa joie de vivre, son enthousiasme et sa si grande sensibilité. Merci Nesrine pour ta complicité, pour ces moments si agréables que j’ai pu partager avec toi (et pour tous ces petits souvenirs que tu nous ramenait de Tunisie!!!). Nos deux voyages en Islande et à Xi’An ainsi que les moments que nous y avons passés ensemble (ainsi que tes achats compulsifs!!) sont gravés pour toujours dans ma mémoire. Je n’oublie pas non plus les membres de l’équipe arrivés plus tard : Thomas Jansson, Irina Poltavchenko, Valentin Barrière, Soumia Dermouche, Béatrice Biancardi.

Comment ne pourrais-je citer ma famille dans ce texte de remerciements. Du fond de mon cœur, je remercie mes parents pour les encouragements qu’ils ont témoigné tout au long de ma – longue – scolarité. Merci de m’avoir transmis le goût d’apprendre et de me dépasser. Merci de m’avoir fait confiance et de m’avoir laissé explorer les chemins qui me tenaient à cœur. Merci de m’avoir soutenu dans les moments les plus difficiles. Je remercie également mes frères et sœurs (particulièrement Thomas, Claire et Lucie pour leur complicité, leur sens de l’humour et du partage). Je me dois également de citer mes cousines (Anne, Cécile et Laura), mes oncles

---

et tantes (Stéphane, Nezha, Marco, Nathalie et Denis) et enfin mes deux grands-mères, Mamie Marcelline et Mamie Nelly, merci pour l'amour que vous avez toujours manifesté.

Je souhaiterais également remercier ceux qui sont un peu ma seconde famille. Tout d'abord, les enfants de Nicolas, Lazare, Marius, Nino. Il m'est impossible d'exprimer verbalement tout ce que vous m'avez apporté et m'apportez encore, vous êtes des enfants formidables et avoir pu partager avec vous un peu de ma vie est une chance que j'accueille tous les jours avec gratitude. Ensuite, Jocelyne, Delphine, Corinne (la team des *ine*), merci de votre présence et de votre soutien pendant ces trois dernières années.

Je tiens également à remercier Antonin, pour son soutien ces deux dernières années. Nous nous sommes rencontrés dans une période difficile pour nous deux. Merci d'avoir été à mes côtés et de m'avoir aidé dans les moments difficiles, mais aussi d'avoir été présent pour partager de nombreux moments de joie.

Enfin, je souhaite dédier l'ensemble de mon travail de thèse à la mémoire de Nicolas Classeau, mon conjoint, mort au Bataclan le 13 novembre 2015. C'est grâce à lui que j'ai décidé de m'inscrire en thèse. Je ne compte pas les soirées qu'il a passé à m'encourager et à me remonter le moral quand ma confiance en moi faisait défaut. Merci pour ta patience quand je devais terminer un article, merci pour les tuyaux en programmation. En te perdant ce soir-là, j'ai perdu un compagnon de route incroyable et je regrette que tu ne puisses être à mes côtés pour voir la concrétisation de ce travail.



# Table des matières

<b>1</b>	<b>Introduction</b>	<b>1</b>
1.1	Analyse de sentiments et conversations humain-agent : questions de recherche . . .	3
1.1.1	Quel phénomène modéliser? . . . . .	3
1.1.2	Comment modéliser? . . . . .	4
1.1.3	Positionnement . . . . .	4
1.2	Spécificités interactionnelles des corpus d'étude . . . . .	5
1.2.1	Présentation des corpus : du <i>small-Talk</i> à la négociation . . . . .	5
1.2.2	Symétrie/complémentarité, coopération/compétition . . . . .	6
1.2.3	Cadre formel vs cadre informel . . . . .	7
1.2.4	Expressions d'attitude et caractéristiques interactionnelles . . . . .	7
1.3	Organisation du mémoire . . . . .	7
<b>I</b>	<b>Etat de l'Art</b>	<b>9</b>
<b>2</b>	<b>Détection d'opinions et analyse automatique de sentiments – Etat de l'Art</b>	<b>11</b>
2.1	Fondements théoriques de l'analyse automatique de sentiments . . . . .	12
2.1.1	Termes et modèles de référence de l'analyse de sentiments . . . . .	12
2.1.2	Le Modèle OCC (Ortony et al., 1990) . . . . .	13
2.1.3	<i>Appraisal in English</i> (Martin and White, 2005) . . . . .	14
2.1.4	Bilan . . . . .	16
2.2	Méthodes en analyse de sentiments . . . . .	17
2.2.1	Aperçu des objectifs et méthodes en analyse de sentiments . . . . .	17
2.2.2	Compositionnalité et polarité . . . . .	18
2.2.3	Approches <i>aspect-based</i> : vers une meilleure analyse des cibles . . . . .	22
2.2.4	Bilan . . . . .	22
2.3	Positionnement : analyse de sentiments pour les interactions humain-agent conversationnel animé . . . . .	23
2.3.1	Choix du modèle théorique . . . . .	23
2.3.2	Choix de la méthode . . . . .	24
<b>II</b>	<b>Fondements linguistiques : les expressions d'attitude dans les conversations</b>	<b>25</b>
<b>3</b>	<b>Les attitudes à l'échelle du tour de parole de l'utilisateur</b>	<b>29</b>
3.1	Description syntaxique et sémantique des attitudes dans la littérature . . . . .	31
3.1.1	Patrons syntaxiques . . . . .	31
3.1.2	Règles sémantiques . . . . .	31
3.2	Formes syntaxiques, rôles sémantiques et calcul de la polarité . . . . .	34
3.2.1	Structures syntaxiques et identification de la source et de la cible . . . . .	34

3.2.2	Structures syntaxiques et règles de polarité . . . . .	38
3.3	Spécificités liées au contexte conversationnel . . . . .	41
3.4	Conclusion du chapitre 3 . . . . .	43
<b>4</b>	<b>Les attitudes à l'échelle de la paire adjacente</b>	<b>45</b>
4.1	Les attitudes dans le cadre collaboratif du dialogue . . . . .	47
4.1.1	Dialogue et contributions . . . . .	47
4.1.2	Les Attitudes, des Actes Participatifs . . . . .	48
4.1.3	Attitudes et Contributions dans les Conversations Humain-Agent . . . . .	49
4.2	Expressions d'attitude strictement collaboratives . . . . .	49
4.2.1	Réponses aux questions fermées . . . . .	50
4.2.2	Réponses aux questions ouvertes . . . . .	51
4.2.3	Assertion d'une attitude . . . . .	53
4.3	Formalisation des attitudes strictement collaboratives . . . . .	54
4.4	<i>Small-Talk</i> vs négociation . . . . .	54
4.5	Conclusion du Chapitre 4 . . . . .	56
<b>5</b>	<b>Les attitudes dans la séquence thématique</b>	<b>57</b>
5.1	Fondements théoriques à la notion de thème . . . . .	59
5.1.1	De la perspective du contenu à celle du fonctionnement . . . . .	59
5.1.2	Le thème dans les conversations humain-agent . . . . .	59
5.2	Attitudes et thèmes dans des conversations humain-agent . . . . .	60
5.2.1	Séquence thématique : <i>small-talk</i> vs négociation . . . . .	60
5.2.2	Vue d'ensemble des thèmes dans le corpus semaine . . . . .	61
5.2.3	Relations sémantiques entre les cibles et les thèmes du scénario Semaine . . . . .	62
5.3	Exemple sur une session du corpus Semaine . . . . .	62
5.4	Conclusion du chapitre 5 . . . . .	63
<b>III</b>	<b>Détection des <i>likes</i> et <i>Dislikes</i> : de la détection à l'évaluation</b>	<b>65</b>
<b>6</b>	<b>Détection des <i>likes</i> et <i>dislikes</i> : du tour de parole à la séquence thématique</b>	<b>69</b>
6.1	Méthodologie et objectifs applicatifs . . . . .	72
6.1.1	Méthodologie : une approche <i>corpus-driven</i> . . . . .	72
6.1.2	Des attitudes aux <i>likes</i> et <i>dislikes</i> . . . . .	74
6.2	Tour de parole de l'utilisateur : détection des expressions simples de <i>likes</i> et <i>dislikes</i>	76
6.2.1	Analyse lexicale . . . . .	77
6.2.2	Analyse syntagmatique . . . . .	79
6.2.3	Analyse phrastique . . . . .	81
6.3	Paire adjacente : détection des <i>likes</i> et <i>dislikes</i> strictement collaboratifs . . . . .	83
6.3.1	Tour de parole de l'agent . . . . .	83
6.3.2	Tour de parole de l'utilisateur . . . . .	84
6.3.3	Calcul des propriétés sémantiques . . . . .	84
6.4	Séquence thématique : optimiser le traitement des cibles . . . . .	86
6.4.1	Niveau A : relations entre les thèmes et les domaines de <i>WordNet-Domains</i>	86
6.4.2	Niveau B : détection de mots-thèmes . . . . .	87
6.4.3	Niveau C : vers un modèle utilisateur . . . . .	87
6.4.4	Exemple sur une séquence du corpus Semaine . . . . .	90
6.5	Conclusion du chapitre 6 . . . . .	92

---

<b>7</b>	<b>Annotation de corpus et protocoles d'évaluation</b>	<b>93</b>
7.1	Évaluation dans le cadre d'un scénario de <i>small-talk</i> (corpus Semaine) . . . . .	96
7.1.1	Évaluation <i>small-talk</i> 1 : tour de parole et paire adjacente . . . . .	96
7.1.2	Évaluation qualitative <i>small-talk topic</i> : séquence thématique . . . . .	100
7.2	Évaluation dans le cadre d'un scénario de négociation . . . . .	102
7.2.1	Protocole et plateforme d'annotation . . . . .	102
7.2.2	Accord inter-annotateur . . . . .	102
7.2.3	Évaluation et discussion . . . . .	103
7.3	Conclusion du chapitre 7 . . . . .	105
<b>8</b>	<b>Conclusion et perspectives</b>	<b>107</b>
8.1	Résumé . . . . .	108
8.2	Contributions . . . . .	109
8.2.1	Formalisation des expressions d'attitude en contexte conversationnel . . .	109
8.2.2	Implémentation et validation des règles linguistiques . . . . .	109
8.2.3	Méthodologie d'évaluation . . . . .	110
8.3	Perspectives . . . . .	110
8.3.1	Amélioration des performances du système . . . . .	110
8.3.2	Évaluation dans le cadre d'une interaction humain-agent . . . . .	111
8.3.3	Vers la gestion des préférences de l'utilisateur . . . . .	111
8.3.4	Vers l'hybridation . . . . .	112
	<b>Bibliographie</b>	<b>120</b>
	<b>Appendices</b>	<b>121</b>
<b>A</b>	<b>Glossaire</b>	<b>123</b>
<b>B</b>	<b>Plateforme pour l'annotation du corpus Semaine</b>	<b>125</b>
<b>C</b>	<b>Jeu d'annotation : <i>likes</i> et <i>dislikes</i> dans le corpus Semaine</b>	<b>129</b>
<b>D</b>	<b>A Web-Based Platform for Annotating Sentiment Related Phenomena in Human-Agent Conversations</b>	<b>133</b>
D.1	Introduction . . . . .	133
D.2	A Platform Focusing on Verbal Content and Conversational Structure . . . . .	133
D.3	Annotating User's Sentiments and Preferences . . . . .	134
<b>E</b>	<b>Implémentation des grammaires dans <i>Unitex</i></b>	<b>137</b>
E.1	Les grammaires hors-contexte . . . . .	137
E.2	Grammaires et graphes dans <i>Unitex</i> . . . . .	137
E.3	Cascade de transducteurs . . . . .	138
<b>F</b>	<b>Détail des grammaires</b>	<b>141</b>
F.1	Grammaires de niveau syntagmatique . . . . .	142
F.2	Grammaires de niveau phrastique . . . . .	147

# Liste des tableaux

2.1	Règles pour l'attribution d'une polarité aux groupes nominaux (Shaikh et al., 2007)	18
2.2	Règles pour l'attribution d'une polarité aux triplets (Shaikh et al., 2007)	19
2.3	Calcul de la polarité au niveau du groupe nominal restreint (Klenner et al., 2009)	20
2.4	Calcul de la polarité au niveau du groupe nominal étendu (Klenner et al., 2009)	20
2.5	Règles pour le calcul de la polarité (Neviarouskaya et al., 2010a)	21
3.1	Synthèse des patrons syntaxiques présentés dans (Bednarek, 2009)	32
3.2	Règles sémantiques de calcul de polarité (Neviarouskaya et al., 2010b)	33
4.1	Exemple d'assertion et de compliment dans des paires adjacentes	50
4.2	Orientation de polarité dans les questions ouvertes	51
4.3	Accord et désaccord à une question fermée	51
4.4	Questions fermées portant sur différents types d'attitude	52
4.5	Questions ouvertes portant sur une caractéristique fondamentale	52
4.6	Questions ouvertes portant sur une caractéristique circonstancielle	52
4.7	Accord et désaccord portant sur des assertions d'attitude	53
6.1	Synthèse des indices lexicaux annotés	78
6.2	Ensemble des groupes syntaxiques traités	80
6.3	Structures syntaxiques des <i>likes</i> et <i>dislikes</i> (tour de parole de l'utilisateur)	82
6.4	Structures syntaxiques et valeurs sémantiques des <i>likes</i> et <i>dislikes</i> (tour de parole de l'agent)	84
6.5	Calcul de la valeur sémantique de la réponse de l'utilisateur	85
6.6	Calcul de la valeur sémantique des attributs <i>polarité</i> , <i>cible</i> et <i>source</i> de la <i>L&amp;DSC</i> détectée	85
7.1	Scores du kappa de Fleiss et de l'alpha de Cronbach obtenus sur les 60 sous-ensembles (évaluation <i>small-talk</i> 1)	97
7.2	Mesures d'accord entre le système et la référence (évaluation <i>Small-Talk</i> 1)	99
7.3	Moyennes des mesures obtenues pour chaque objet	102
7.4	Mesures d'accord entre le système et la référence (évaluation négociation)	104
E.1	Cascade de grammaires <i>Unitex</i> utilisée par notre système	140
F.1	Liste des symboles utilisés dans les grammaires	141
F.2	Grammaires traitant les syntagmes verbaux intégrant un verbe d'opinion	142
F.3	Grammaire traitant les syntagmes verbaux exprimant un <i>like</i> ou un <i>dislike</i>	143
F.4	Grammaires traitant les syntagmes verbaux attributifs	144
F.5	Grammaires traitant les syntagmes verbaux n'appartenant pas aux catégories <i>attributif</i> , <i>opinion</i> , <i>like</i> ou <i>dislike</i>	145
F.6	Grammaires traitant les syntagmes nominaux, les syntagmes adjectivaux et les déterminants	146
F.7	Grammaire reconnaissant les structures de type <Target> <Make> <Affect> <Source>	147
F.8	Grammaire reconnaissant les structures de type attributif – Partie 1	148

F.9	Grammaire reconnaissant les structures de type attributif – Partie 2 . . . . .	149
F.10	Grammaire reconnaissant les structures de type <Source><Have><Affect> . .	149
F.11	Grammaire reconnaissant les structures de type <Source><Like><Target> . . .	150
F.12	Grammaire reconnaissant les structures de type <Source><Feel><Affect><Target>	150
F.13	Grammaire reconnaissant les structures à verbes d’opinion . . . . .	151

# Table des figures

2.1	Structure cognitive des émotions selon le modèle OCC . . . . .	14
2.2	Vue d'ensemble du modèle de Martin and White (2005) . . . . .	16
3.1	Syntagmes adjectivaux et nominaux référant à une attitude . . . . .	34
3.2	Syntagme référant à une attitude en position attribut . . . . .	35
3.3	Syntagme référant à une attitude en position sujet . . . . .	35
3.4	Structure attributive à verbe d'opinion avec ellipse du verbe attributif . . . . .	37
3.5	Structure attributive à verbe d'opinion sans ellipse du verbe attributif . . . . .	37
3.6	Structure à prédicat nominal . . . . .	38
3.7	Structures à prédicat verbal . . . . .	38
3.8	Règles de polarité dans le syntagme adjectival . . . . .	39
3.9	Calcul de la polarité dans le syntagme nominal . . . . .	39
3.10	Inversion de polarité dans le syntagme verbal . . . . .	39
3.11	Inversion de polarité dans le cadre d'une structure attributive . . . . .	40
3.12	Application de la règle de neutralisation dans le cadre d'une structure à verbe d'opinion . . . . .	40
3.13	Attitudes dans le tour de parole de l'utilisateur. Extrait du corpus Semaine . . . . .	42
3.14	Attitudes dans le tour de parole de l'utilisateur. Extrait du corpus de négociation . . . . .	42
4.1	Typologie des paires adjacentes (Clark and Schaefer, 1989) . . . . .	48
4.2	Expressions d'attitude strictement collaboratives. Extrait du corpus Semaine . . . . .	55
4.3	Expressions d'attitude strictement collaboratives. Extrait du corpus de négociation . . . . .	55
5.1	Relation sémantique entre la cible et le thème . . . . .	61
5.2	Structure thématique dans une session du corpus Semaine . . . . .	63
6.1	Schéma d'annotation . . . . .	73
6.2	<i>Balance Theory</i> de Heider . . . . .	74
6.3	Traitement de expressions simples de <i>likes</i> et <i>dislikes</i> . . . . .	76
6.4	Analyse lexicale du tour de parole de l'utilisateur . . . . .	77
6.5	Analyse syntagmatique dans le tour de parole de l'utilisateur . . . . .	79
6.6	Analyse phrastique du tour de parole de l'utilisateur . . . . .	81
6.7	Liens entre les thèmes identifiés dans le corpus <i>Semaine</i> et les domaines de <i>WordNet-Domain</i> . . . . .	87
6.8	Sorties des niveaux A et B . . . . .	90
6.9	Analyses du niveau C . . . . .	91
7.1	Processus d'annotation de la campagne Semaine . . . . .	98
B.1	Questionnaire 1 de la plateforme Semaine . . . . .	125
B.2	Questionnaire 2 de la plateforme Semaine . . . . .	126
B.3	Questionnaire 3 de la plateforme Semaine . . . . .	126
B.4	Questionnaire 4 de la plateforme Semaine . . . . .	127
B.5	Questionnaire 5 de la plateforme Semaine . . . . .	127

B.6	Questionnaire 6 de la plateforme Semaine . . . . .	128
C.1	Informations relatives aux paires (jeu d’annotations du corpus Semaine) . . . . .	130
C.2	Nombre d’expressions de <i>like</i> ou <i>dislike</i> pour chaque paire (jeu d’annotations du corpus Semaine) . . . . .	131
C.3	Caractéristiques des expressions de <i>like</i> et <i>dislike</i> (jeu d’annotations du corpus Semaine) . . . . .	132
D.1	Platform Overview . . . . .	134
D.2	First questionnaire of Task 1 . . . . .	135
D.3	First questionnaire of Task 1 . . . . .	135
E.1	Exemple de graphe Unitex . . . . .	137
E.2	Graphe <i>Unitex</i> permettant de détecter des expressions de <i>like</i> ou <i>dislike</i> de type “I feel happy about life” . . . . .	139

# Chapitre 1

## Introduction

### Sommaire

---

1.1	Analyse de sentiments et conversations humain-agent : questions de recherche	3
1.1.1	Quel phénomène modéliser? . . . . .	3
1.1.2	Comment modéliser? . . . . .	4
1.1.3	Positionnement . . . . .	4
1.2	Spécificités interactionnelles des corpus d'étude . . . . .	5
1.2.1	Présentation des corpus : du <i>small-Talk</i> à la négociation . . . . .	5
1.2.2	Symétrie/complémentarité, coopération/compétition . . . . .	6
1.2.3	Cadre formel vs cadre informel . . . . .	7
1.2.4	Expressions d'attitude et caractéristiques interactionnelles . . . . .	7
1.3	Organisation du mémoire . . . . .	7

---

**N**OTRE thèse se situe à la croisée de deux domaines de recherche : celui de l’analyse de sentiments (*sentiment analysis* en anglais) et celui des agents conversationnels animés. Les agents conversationnels animés peuvent être définis comme des personnages virtuels, c’est-à-dire créés par un programme informatique, ayant la capacité de converser avec un utilisateur humain. Ce sont “des artefacts virtuels capables de simuler la richesse de la conversation humaine”, tant dans sa dimension verbale que non-verbale (Ochs et al., 2013). Les agents jouent différents rôles – assistant, tuteur, compagnon – et sont impliqués dans différents types d’interaction avec l’utilisateur. Dans cette thèse, les agents conversationnels évoqués sont de forme humanoïde, mais de manière généralement, ils peuvent s’incarner dans des formes plus ou moins réalistes. Afin d’accroître les compétences communicationnelles de l’agent, il est important que celui-ci soit doté d’une forme d’intelligence socio-émotionnelle. Différents travaux ont ainsi montré que ces capacités socio-émotionnelles permettent d’améliorer de manière significative la qualité de l’interaction (Beale and Creed, 2009). L’agent doit ainsi être capable de gérer des signaux socio-émotionnels, tant du côté de la production que de celui de la compréhension. Il doit être en mesure non seulement de comprendre les signaux émotionnels émis par l’utilisateur – expressions faciales, caractéristiques acoustiques, gestes, etc. – mais aussi de produire de tels signaux. Cette gestion de la dimension sociale et émotionnelle doit s’exécuter de manière intelligente : ayant recours à des stratégies d’interaction socio-émotionnelle, l’agent va produire des signaux sociaux et émotionnels, après avoir déterminé la nature de l’émotion à exprimer (Pecune et al., 2014).

En ce qui concerne la génération de nombreux travaux ont produit des modèles optimisant la production de gestes ou d’expressions faciales pour l’expression soit d’émotions soit d’attitudes sociales ((De Rosis et al., 2003), (Chollet et al., 2014)). Les travaux existant cherchent à relever différents challenges : la production d’expressions faciales affectives appropriées lorsque l’agent écoute mais aussi lorsqu’il parle (Ding et al., 2013) ; la génération d’expressions corporelles exprimant des émotions (Fourati and Pelachaud, 2015) ; ou la production de signaux non-verbaux exprimant des attitudes sociales, c’est-à-dire “ affective style[s] that spontaneously develop or [are] strategically employed in the interaction with a person or a group of persons” (Scherer, 2005).

En ce qui concerne la détection des comportements socio-émotionnels en interaction, c’est essentiellement le non verbal qui a jusqu’à présent primé. Composante essentielle de l’interaction, la dimension non-verbale est analysée par la prise en compte de différentes caractéristiques : expressions faciales, *action units* ou indices paralinguistiques et acoustiques (Schuller et al., 2011). Le contenu verbal est quant à lui encore rarement exploité dans les interaction humain-agent. Or, si l’analyse du non-verbal est indispensable, il est important de noter que certains éléments ne peuvent être correctement interprétés que par une analyse détaillée du contenu verbal. Une opinion sur une entité ou une personne pourra difficilement être comprise grâce à l’analyse exclusive du contenu non-verbal.

Cette thèse s’intéresse à l’analyse contenu verbal produit par l’utilisateur et à la manière dont celui-ci réfère ou exprime des sentiments, des affects ou des attitudes<sup>1</sup>. Elle est consacrée au développement d’un modèle de détection de ces expressions au cours d’une interaction multimodale et en face-à-face avec l’agent conversationnel animé. Parmi les travaux antérieurs, deux études seulement présentent un système intégrant un module de détection des sentiments dans le contenu verbal pour les interactions humain-agent (Smith et al., 2011a, Yildirim et al., 2011). Ces méthodes ont le mérite de proposer une première approche du problème. Néanmoins, elles diffèrent peu de ce qui est déjà proposé dans le domaine de l’analyse d’opinion et des sentiments et ne sont pas conçues pour des interactions socio-affectives. Or, si l’intégration d’un système de détection des sentiments dans le cadre d’interaction humain-agent apparaît comme nécessaire, son développement requiert cependant la prise en compte d’un certain nombre de contraintes et d’enjeux propres à ce cadre applicatif.

---

1. La suite de notre introduction, ainsi que l’état de l’art apporte plus précisions quant à cette terminologie

Dans cette introduction, nous présentons dans un premier temps les questions de recherche qui ont guidé notre travail et qui concernent l'adaptation d'un modèle de détection des sentiments au contexte spécifique des conversations humain-agent. Dans un deuxième temps, nous présentons les corpus sur lesquels nous nous sommes appuyée pour développer notre modèle, ainsi que les spécificités du cadre d'interaction qu'ils impliquent. Enfin, dans un troisième temps, nous détaillons l'organisation de notre mémoire.

## 1.1 Analyse de sentiments et conversations humain-agent : questions de recherche

Le contexte applicatif dans lequel se place notre travail, celui des conversations humain-agent, est assez éloigné des contextes plus traditionnellement ciblés en analyse de sentiments. Ainsi, le développement d'un modèle de détection dans un tel contexte soulève un ensemble de problématiques nouvelles auxquelles nous avons dû nous confronter durant notre travail de thèse et que nous synthétisons dans cette section. Le premier concerne la délimitation du type de expressions pouvant aider l'agent dans la gestion de l'interaction. La deuxième concerne le choix de la méthode parmi celles déjà utilisées dans le domaine de l'analyse de sentiments. La troisième concerne la nature conversationnelle des phénomènes à traiter.

### 1.1.1 Quel phénomène modéliser ?

Dans le domaine de l'analyse de sentiments, les systèmes de détection proposés ciblent différents phénomènes. Comme l'explique Clavel and Callejas (2016), tandis que certains travaux se concentrent sur des sous-phénomènes spécifiques en portant une attention particulière aux expressions référant à des comportements affectifs ou émotionnels (Ishizuka, 2012, Perikos and Hatzilygeroudis, 2013, Neviarouskaya et al., 2010a), d'autres englobent, sous l'étiquette de *sentiment*, un nombre important d'expressions variées. Ces derniers ne fournissent pas toujours une description détaillée des expressions et la caractéristique distinctive la plus souvent mise en valeur est la polarité, positive ou négative. Du point de vue de l'agent, si la détection de sentiments positifs et négatifs exprimés verbalement est un atout non négligeable pour la gestion de l'interaction, une catégorisation fine apparaît tout aussi nécessaire. Prenons l'exemple des phrases ci-dessus.

- (i) "This movie is so beautiful"
- (ii) "I'm very happy today"
- (iii) "It is so pleasant to hear that"

Schématiquement, tous ces énoncés peuvent être détectés comme des expressions positives. Néanmoins, il faut s'interroger sur leur intérêt sémantique du point de vue de l'agent. Tous ne véhiculent pas le même type d'information. Tandis que le premier informe sur les opinions de l'utilisateur, le deuxième informe sur son état émotionnel et le troisième met en relation cet état émotionnel avec un déclencheur. Les deux derniers mettent l'accent sur le locuteur, alors que le premier met l'accent sur l'attribution par ce dernier d'une valeur évaluative à l'objet *movie*. Ces énoncés, bien qu'ayant la même polarité, n'appellent pas le même type de réponse et ne peuvent donc être traités de manière similaire. De plus, selon le cadre applicatif, l'agent pourra avoir besoin d'information concernant soit les émotions et les affects, soit les opinions, appréciations et jugements. Une catégorisation plus fine des expressions de sentiment apparaît donc indispensable, une catégorisation offrant des informations précises nécessaires à la sélection adaptée d'une réponse par le moteur de dialogue. Dans la mesure où la catégorisation doit contribuer à l'amélioration de la gestion de l'interaction, il paraît nécessaire que, parmi la vaste classe des

sentiments, on puisse identifier des sous-catégories, et que parmi ces catégories, soient définies celles étant les plus adaptées aux buts communicationnels de l’agent.

### 1.1.2 Comment modéliser ?

Comme cela sera décrit dans le chapitre consacré à l’état de l’art (Chapitre 2), différents types de méthodes ont été utilisés dans le domaine de l’analyse de sentiments : méthodes de classification, méthodes de *deep-learning*, méthodes exploitant des règles expert. Dans le cadre d’un système de détection des sentiments appliqué aux conversations humain-agent, différents aspects sont à prendre en compte pour choisir la méthode la plus adaptée. Tout d’abord, il est nécessaire à s’interroger sur le niveau de précision de l’analyse produite. Cela rejoint ce que nous avons décrit ci-dessus sur la distinction nécessaire entre les différentes sous-catégories d’expressions de sentiment. A cette distinction catégorielle s’ajoute une distinction fondée sur la nature de la source et de la cible. Il semble en effet indispensable de pouvoir s’assurer que les expressions détectées aient bien comme source l’utilisateur, et non un autre individu dont il pourrait rapporter les propos. De même, l’identification de la cible est une phase indispensable pour produire des résultats exploitables par l’agent. Une analyse pertinente doit permettre d’indiquer, par exemple, si la cible est un élément connu de l’agent, un élément que lui-même a pu introduire dans la conversation. Dans l’idéal, elle devrait également définir pour toute expression si la cible est différente de celles déjà identifiées au préalable, ou s’il s’agit d’un nouvel élément. Prenons l’exemple des phrases ci-dessous. L’analyse doit pouvoir définir les cibles *movie* et *Acopalypse Now* comme référent à la même entité. Cet aspect nous amène également à la représentation et la formalisation des résultats.

(iv) “I didn’t like *Acopalypse Now*, I really didn’t like this movie”

Sur un plan plus technique, le choix de la méthode doit reposer sur une connaissance des ressources disponibles. Ainsi, dans le cadre d’une méthode par apprentissage automatique supervisé, un corpus annoté est indispensable pour les phases d’apprentissage et de test. Or, à notre connaissance, aucun corpus conversationnel incluant des annotations de sentiments tels qu’ils sont exprimés dans le contenu verbal n’existe. Cette donnée a du être prise en compte au moment du développement de notre modèle. Enfin, la méthode choisie doit être exploitable dans le cadre d’une conversation en face-à-face et en temps réel. Cela suppose qu’elle puisse traiter une parole de nature conversationnelle. Dans un tel cadre, l’identification d’expressions dans le cadre d’énoncés prononcés par l’utilisateur est indispensable, mais la prise en compte de contexte d’analyse plus large l’est tout en autant. Ainsi, le système développé doit pouvoir s’adapter au *turn-taking* et analyser chaque tour de parole de l’utilisateur ainsi que les relations sémantiques entretenues avec le contenu verbal de l’agent.

### 1.1.3 Positionnement

Pour répondre à ces différentes problématiques, nous avons choisi de développer un modèle de détection exploitant un ensemble de règles linguistiques. Ces règles s’appuient sur le modèle théorique présenté par Martin and White (2005) et que nous décrivons en détail dans l’état de l’art (Chapitre 2). Ce modèle propose une hiérarchie et une description des attitudes (dont la définition diffère de celle proposée par Scherer (2005)), regroupant l’ensemble des expressions verbales référant à des évaluations axiologiques ou à des réactions émotionnelles. A partir des données extraites de la littérature et d’une analyse de conversations humain-agent, nous avons développé des règles permettant de formaliser les expressions d’attitudes telles qu’elles se manifestent dans un contexte de conversation humain-agent. A l’intérieur de la conversation, différentes unités ont été considérées : le tour de parole, la paire adjacente et la séquence thématique. Les règles, telles qu’elles ont été implémentées (voir Chapitre 6), concernent pour chacune d’elles une des trois unités spécifiques. Pour ce travail, nous avons fait le choix de ne

pas traiter les indices multi-modaux. Nous considérons en effet qu'il est nécessaire de développer des règles se concentrant sur le contenu verbal avant de pouvoir envisager la manière de traiter conjointement indices verbaux et non-verbaux.

### 1.2 Spécificités interactionnelles des corpus d'étude

Le modèle de détection présenté dans cette thèse est fondé sur l'étude d'un certain type de conversations humain-agent. Ces conversations sont en face à face, l'échange est multimodal et se fait de parole à parole et enfin la langue utilisée est l'anglais. Deux types d'interaction humain-agent ont été au cœur de notre travail. Tout d'abord, un modèle d'interaction de type *small-talk*, où l'utilisateur et l'agent discutent de sujets de la vie quotidienne, ensuite, un modèle d'interaction de type négociation où l'agent et l'utilisateur débattent et négocient pour parvenir à un accord, à une entente. Pour comprendre les caractéristiques de ces deux types d'interaction, nous reprenons ici la typologie des interactions verbales dressée par Vion (1992) pour caractériser les scénarios .

#### 1.2.1 Présentation des corpus : du *small-Talk* à la négociation

Afin de mener les observations nécessaires à la création des règles linguistiques qui constituent notre système, nous avons exploité deux corpus : le corpus Semaine McKeown et al. (2011) et un corpus de négociation développé par Gratch et al. (2016).

Le corpus Semaine McKeown et al. (2011) est composé de 65 transcriptions manuelles de sessions où un utilisateur humain interagit avec un opérateur humain jouant le rôle d'un agent conversationnel. Ces interactions sont fondées sur un scénario impliquant quatre personnages d'agent émotionnellement typés : Poppy, joyeuse et extravertie, Prudence, sensible et raisonnable, Spike, colérique et conflictuel, et enfin Obadiah, dépressif et maussade. Leur objectif est de pousser l'utilisateur vers l'état émotionnel qu'ils incarnent. Les énoncés que peuvent prononcer les opérateurs sont contraints par un script – néanmoins, certains énoncés dans le corpus s'écartent du script. Parmi les 65 sessions à notre disposition, nous avons utilisé 30 sessions pour le développement et 35 pour l'évaluation (voir le chapitre 7).

Pour la négociation, nous avons utilisé le corpus de Gratch et al. (2016) composé de 154 sessions transcrites manuellement, opposant deux participants. Selon le protocole, les joueurs doivent trouver un accord afin de se partager le contenu d'un casier. Celui est composé d'objets de valeur appartenant à trois types différents (3 caisses de disques, 2 lampes et 1 peinture, pour le premier ensemble, et 3 caisse d'assiettes, 2 chaises et 1 horloge, pour le second). Chaque joueur reçoit une récompense pour chaque accord possible, généralement inconnue de l'autre partie. Dans le scénario, deux structures de négociation sont possibles en fonction des intérêts partagés par les joueurs. La structure intégrative concerne une négociation de somme non-nulle, gagnant-gagnant, alors que la structure distributive implique une négociation compétitive, à somme nulle. Les joueurs ne connaissent pas à l'avance les intérêts de leurs adversaires. Ce corpus comporte deux parties : la première implique des négociations humain-humain, la seconde une négociation entre des utilisateur humain et un agent. Notre travail se concentre uniquement sur les négociations humain-agent collectées à l'aide d'un protocole *wizard-of-oz*. Nous avons utilisé 77 pour le développement du système et les 77 autres pour l'évaluation (voir la session 7).

Le corpus Semaine et le corpus de négociation reposent sur des scénarios d'interaction différents. Les objectifs communicationnels de l'agent ne sont pas les mêmes et la manière dont l'utilisateur va pouvoir interagir avec lui en est nécessairement affecté. Pour comprendre la nature du cadre d'interaction et la place attribuée à l'utilisateur dans chacun de ces scénarios, nous proposons dans la section suivante de les caractériser en nous appuyant sur la typologie proposée par R. Vion, dans son ouvrage (Vion, 1992).

### 1.2.2 Symétrie/complémentarité, coopération/compétition

Dans son ouvrage (Vion, 1992), R. Vion prend en considération quatre critères qui s’opposent de manière binaire et qui lui permettent de dresser une typologie des interactions.

- **Symétrie/Complémentarité** : “Dans le premier cas, les partenaires ont tendance à adopter un comportement en miroir, leur interaction peut donc être dite symétrique. [...] Dans le second cas, le comportement de l’un des partenaires complète celui de l’autre. Une interaction symétrique se caractérise donc par l’égalité et la minimisation de la différence, tandis qu’une interaction complémentaire se fonde sur la maximalisation de la différence”. Ces deux concepts s’opposent dans leur définition, néanmoins, tels qu’ils s’incarnent dans les interactions, ils ne sont pas exclusifs l’un de l’autre. Comme l’explique Vion, une interaction donnée pourra fonctionner sur le principe de complémentarité, mais également inclure des moments conversationnels faisant davantage appel au principe de symétrie.
- **Coopération/Compétitivité** : La coopération concerne des interactions où le modèle d’activité est non stratégique et orienté vers l’intercompréhension, tandis que la compétitivité concerne des interactions où le modèle d’activité est stratégique et orienté vers le succès. La conversation représente le type même d’interaction de nature coopérative, tandis que la dispute ou le débat relèvent davantage de la compétition.

Les critères de coopération/compétitivité et de complémentarité/symétrie nous offrent la possibilité de poser un éclairage intéressant sur les interactions humain-agent que nous avons à traiter. Comme cela a déjà été précisé, dans les conversations humain-agent, c’est le scénario qui définit le cadre d’interaction. Concernant les interactions de négociation, le cadre est clairement compétitif. Dans ce type d’interaction, le scénario contraint l’agent à définir des stratégies en vue d’obtenir quelque chose. L’utilisateur, s’il joue le jeu, doit entretenir le même rapport et s’engager dans l’interaction afin de créer les conditions de son succès ou de son échec. Cela n’est pas le cas dans une conversation évoquant des sujets de la vie quotidienne où l’objectif est davantage de comprendre et de se faire comprendre. Le cadre est coopératif. L’agent est en recherche d’informations sur l’utilisateur, son objectif de communication ne le place donc pas en compétition avec l’utilisateur. Ce dernier, s’il accepte l’interaction, cherchera à se faire comprendre de l’agent. Bien sûr, comme cela a été dit, il sera possible d’y trouver des moments où l’utilisateur s’inscrit dans un rapport de force avec l’agent en manifestant une volonté de débattre par exemple, mais la tendance globale de l’interaction restera orientée vers l’intercompréhension.

Le critère de symétrie apparaît comme un critère valable pour les interactions de négociation. Si l’utilisateur souhaite parvenir, au terme de la négociation, à un succès, il a tout intérêt à adopter un comportement miroir à l’agent et de négocier avec le même degré d’implication. Cela ne suppose pas que ses stratégies de négociation seront identiques à celles de l’agent, mais qu’il ne maximisera pas sa différence sur le plan verbal et dialogique. Pour les conversations de type *small-talk*, la symétrie pourrait, à la manière des conversations humain-humain, être considérée comme l’un des critères définitoires du cadre interactif. Néanmoins, leur caractère non-spontané et le fait que la progression de la conversation soit contrainte par le scénario et les thèmes sélectionnés par le moteur de dialogue, encourage à les placer du côté de la complémentarité. R. Vion explique que “dans une relation complémentaire, il y a deux positions différentes possibles [pour chacun des locuteurs]. L’un des partenaires occupe une position qui a été désignée comme supérieure, première ou “haute” (*one-up*), et l’autre la position correspondante dite inférieure, seconde ou “basse” (*one-down*)”. Dans le scénario Semaine, l’agent occupe, dans une certaine mesure, une position haute puisque lui seul peut poser des questions et donner à la conversation son orientation. Ces conversations peuvent donc être décrites comme des interactions coopératives où l’agent et l’utilisateur jouent des rôles complémentaires.

### 1.2.3 Cadre formel vs cadre informel

Selon R. Vion, le caractère formel/non-formel de l'interaction dépend de plusieurs autres facteurs : le nombre de participants ; la nature du cadre interactif ; les règles de circulation de la parole ; le caractère ouvert ou fermé du contrat de parole ; la nature des rituels d'ouverture et de clôture ; la nature des signaux d'écoute. Tant pour la négociation que pour le *small-talk*, nous pouvons considérer que nous sommes plutôt dans un cadre qui tend vers le formel. Même si les interactions de type *small-talk* tendent à imiter les conversations spontanées humain-humain, elles ne le sont pas réellement – du moins dans les corpus que nous avons pu étudier. Un protocole définit la manière dont l'utilisateur peut interagir avec l'agent. Le caractère virtuel de l'agent place de fait l'utilisateur dans un cadre non spontané.

### 1.2.4 Expressions d'attitude et caractéristiques interactionnelles

Les caractéristiques décrites ci-dessus constituent une contrainte qui conditionne la nature du contenu verbal produit par l'utilisateur. Ainsi, dans une conversation de type *small-talk*, la complémentarité contraint l'utilisateur à une position *one-down*. Cela suppose que l'utilisateur suive le cadre posé par l'agent. Dans le cas contraire, l'interaction ne pourra pas fonctionner. De même, dans le cadre d'une négociation, l'utilisateur doit entrer dans ce rapport de compétitivité avec l'agent, accepter de jouer le jeu de la négociation.

Dans la mesure où, pour permettre à l'interaction de fonctionner, l'utilisateur doit respecter ces contraintes, nous pouvons considérer que les attitudes qu'il exprimera ne le seront qu'à l'intérieur de ce cadre. Puisque des écarts peuvent néanmoins être envisagés, nous avons choisi de n'intégrer à notre modèle que les attitudes respectant ce cadre. Le modèle ne concerne donc pas les attitudes n'ayant pas de contenu à la conversation. Sont ainsi exclues les attitudes ne s'adressant pas à l'agent, celles dont la cible est un élément physique du dispositif d'interaction (installation audio-vidéo, l'agent lui-même) ou bien l'interaction elle-même (par exemple, tout commentaire portant sur l'intérêt/désintérêt de la conversation). De plus, le modèle doit pouvoir prendre en compte que selon le type de scénario traité les attitudes exprimées puissent varier dans leur forme. Pour cela, nous avons cherché à développer un modèle ayant un socle suffisamment générique pour être applicable à différents scénarios, mais qui intègre également leurs spécificités.

## 1.3 Organisation du mémoire

Le mémoire s'organise en trois parties. La première partie est consacrée à l'état de l'art, la deuxième à la description et à la formalisation du phénomène des attitudes dans le contexte des conversations humain-agent, et la troisième à l'architecture du système de détection et aux différentes évaluations que nous avons menées.

- La **partie I** explore les différentes méthodes et applications utilisées en analyse de sentiments. Elle est pour nous l'occasion d'affirmer nos choix théoriques et techniques. Après une présentation des fondements et des modèles théoriques utilisés dans le domaine, nous abordons la question des méthodes. L'accent est mis sur les travaux favorisant une analyse détaillée des expressions et exploitant soit des méthodes à base de règles expert, soit des méthodes hybrides associant règles expert et algorithme d'apprentissage automatique.
- La **partie II** est consacrée aux fondements théoriques du système que nous avons développé. Elle se compose de trois chapitres offrant pour chacun d'eux un éclairage différent sur la question des attitudes (Martin and White, 2005) telles qu'elles se manifestent en conversation dans le contenu verbal de l'utilisateur.

- Le **chapitre 3** est consacré à la description des attitudes dans le cadre du tour de parole. Sur la base des données extraites de la littérature et des deux corpus d'étude, nous explorons la structure syntaxique et sémantique des énoncés référant à des attitudes et proposons une formalisation sous forme de patrons.
- Le **chapitre 4** s'attache à la description des expressions d'attitude dans le cadre de la paire adjacente. Après une caractérisation des attitudes comme des actes participatifs contribuant à la progression de la conversation et aux connaissances partagées élaborées pendant l'interaction, nous abordons la question des attitudes que nous nommons « strictement collaboratives » et en proposons une description ainsi qu'une formalisation sémantique.
- Enfin, le **chapitre 5** a un cadre d'étude plus large, celui de la séquence thématique. En analysant la manière dont s'articulent les thèmes dans les conversations humain-agent, nous proposons une modélisation du lien sémantique existant entre la cible et le thème.
- La **partie III** est consacrée à la description du système et aux différentes évaluations que nous avons menées.
  - Dans le **chapitre 6**, nous définissons un phénomène spécifique, celui des *likes* et des *dislikes*, que nous considérons comme une sous-catégorie transverse à la notion d'attitudes. Puis, dans un second temps, nous décrivons l'architecture générale du système et les différentes règles exploitées aux trois niveaux d'analyse déjà utilisés dans la partie II : tour de parole, paire adjacente et séquence thématique.
  - Le **chapitre 7** suit la même structure et décrit les trois évaluations menées au cours de notre thèse. La première s'attache davantage aux performances du système pour la détection des *likes* et des *dislikes* dans le cadre du tour de parole et de la paire adjacente, dans le corpus Semaine. La deuxième s'intéresse aux performances du système toujours dans le cadre du tour de parole et de la paire adjacente mais sur le corpus de négociation. Enfin, la troisième évalue les traitements effectués dans le cadre de la séquence thématique. Les deux premières sont des évaluations quantitatives ayant recours à deux corpus annotés et la dernière se présente davantage comme une évaluation qualitative. Ce chapitre est également pour nous l'occasion de décrire les deux campagnes d'annotation menées pour construire nos corpus de référence pour les deux premières évaluations.

**Première partie**

**Etat de l'Art**



## Chapitre 2

# Détection d'opinions et analyse automatique de sentiments – Etat de l'Art

### Sommaire

---

2.1	Fondements théoriques de l'analyse automatique de sentiments . . . . .	12
2.1.1	Termes et modèles de référence de l'analyse de sentiments . . . . .	12
2.1.2	Le Modèle OCC (Ortony et al., 1990) . . . . .	13
2.1.3	<i>Appraisal in English</i> (Martin and White, 2005) . . . . .	14
2.1.4	Bilan . . . . .	16
2.2	Méthodes en analyse de sentiments . . . . .	17
2.2.1	Aperçu des objectifs et méthodes en analyse de sentiments . . . . .	17
2.2.2	Compositionnalité et polarité . . . . .	18
2.2.3	Approches <i>aspect-based</i> : vers une meilleure analyse des cibles . . . . .	22
2.2.4	Bilan . . . . .	22
2.3	Positionnement : analyse de sentiments pour les interactions humain-agent conversationnel animé . . . . .	23
2.3.1	Choix du modèle théorique . . . . .	23
2.3.2	Choix de la méthode . . . . .	24

---

L’ANALYSE de sentiments (ou *sentiment analysis* en anglais) a connu depuis le début des années 2000 de nombreux développements et de nombreux succès tant dans le domaine de la recherche académique que dans celui de l’industrie. Ce succès est favorisé par le développement massif de données issues du web et particulièrement des réseaux sociaux où l’utilisateur est amené à donner son avis sur divers objets. Ceci a entraîné la production de données affectives et axiologiques posant, de par leur variabilité, la question de leur analyse et de leur exploitation. Cette subjectivité évaluative prend diverses formes : elle peut s’exprimer textuellement, dans les forums, les commentaires, les tweets ; elle peut s’ancrer dans un contexte multimodal, dans des vidéos ou être accompagnée de photos. Cette variabilité conduit les chercheurs à développer et à évaluer différents modèles exploitant des technologies issues de l’extraction d’information ou de la fouille de données.

Notre champ d’application diffère de celui traditionnellement exploité en *sentiment analysis*. Néanmoins, une étude des différents modèles et méthodes utilisés dans ce domaine est une étape nécessaire pour définir nos propres choix et positionner notre approche. Notre objectif ici n’est pas de dresser une liste exhaustive des travaux réalisés en *sentiment analysis*, mais de mettre en lumière différentes tendances, afin de définir les outils théoriques et opérationnels pertinents sur lesquels se fonde le développement de notre modèle. Ce chapitre s’intéresse dans un premier temps aux modèles théoriques permettant de décrire la notion de sentiment ou d’opinion. Dans un second temps, nous présentons différents types de méthodes utilisées en *sentiment analysis* et mettons en relief leurs objectifs de développement (classification, extraction d’aspects, calcul de polarité). Après avoir mis l’accent sur les objectifs que se fixe aujourd’hui le *sentiment analysis*, nous présentons deux types d’approches qui se distinguent, selon nous, par leur manière d’envisager la complexité des expressions verbales de sentiments. Enfin, dans un troisième temps, après une rapide présentation de rares travaux intégrant un module de *sentiment analysis* dans une architecture de dialogue humain-machine, nous présentons notre positionnement.

## 2.1 Fondements théoriques de l’analyse automatique de sentiments

Dans le domaine du *sentiment analysis* (SA), différents termes sont utilisés pour référer aux phénomènes linguistiques ciblés : *sentiments, opinions, jugements, affects, attitudes, appréciations, émotions*. Leur définition varie d’une étude à une autre, de même que le type de d’expressions verbales qu’ils permettent de circonscrire. Ce chapitre a pour objet de faire le point sur les fondements théoriques exploités en SA. Dans un premier temps (Sous-section 2.1.1), nous passons en revue les termes et modèles relatifs à la question des sentiments, telle qu’elle est posée dans le domaine, et interrogeons la manière dont sont circonscrites les expressions candidates à la détection. Dans un deuxième et troisième temps (Sous-sections 2.1.2 et 2.1.3), nous interrogeons les domaines de la psychologie et de la linguistique pour nous intéresser à deux modèles fréquemment cités dans les travaux de *sentiment analysis* : le modèle OCC (Ortony et al., 1990) et le modèle de Martin and White (2005).

### 2.1.1 Termes et modèles de référence de l’analyse de sentiments

Définir le phénomène cible apparaît comme un pré-requis nécessaire au développement d’un système d’analyse de sentiments. Cette question émerge dans les premiers travaux produits dans le domaine par la référence à la notion de subjectivité. Dans (Wiebe, 2000), la subjectivité est envisagée comme la caractéristique discriminante des expressions d’opinion et de sentiment. Deux catégories sont envisagées : celle des expressions subjectives, qui portent la marque de ce que pense ou aime le locuteur, et celle des expressions objectives qui exposent un fait sans que ne se manifeste l’avis ou les goûts du locuteur. Les premiers travaux de classification se proposent ainsi de détecter automatiquement des adjectifs (Wiebe, 2000) et des phrases subjec-

tives ((Riloff and Wiebe, 2003), (Wiebe and Riloff, 2005)). A cette notion, s'ajoute celle de la valence, utilisée pour distinguer les expressions subjectives ayant une polarité positive de celles ayant une polarité négative. Dans (Taboada et al., 2011), la polarité est définie relativement au modèle d'Osgood "It usually captures an evaluative factor (positive or negative) and potency or strength (degree to which the word, phrase, sentence or document in question is positive or negative) towards a subject topic, person or idea". Néanmoins, force est de constater que peu d'études réfèrent à un modèle spécifique pour définir cette notion de valence ou de polarité.

Afin de dépasser ces oppositions subjectif vs. objectif et positif vs. négatif et de proposer une approche plus catégorielle, certains travaux se sont efforcés de référer à des modèles théoriques offrant une description plus complexe de ces phénomènes. Dans la vaste classe des expressions de sentiments, les sous-classes d'*affect* et d'*émotion* retiennent l'attention de quelques chercheurs. Le travail de Neviarouskaya et al. (2010a) exploite le modèle d'Izard (Izard, 1977) pour définir les neuf classes d'affects que le système est capable de détecter (*colère, dégoût, peur, honte, intérêt, joie, tristesse, surprise, culpabilité*). Une démarche similaire peut être observée par Perikos and Hatzilygeroudis (2013), qui retiennent les six classes d'émotions basiques décrites par Ekman (1999). Toujours dans un objectif de catégorisation fine des phénomènes, Wiebe et al. (2005) développent un schéma d'annotation *private states* tels que définis par Quirk (1985). Ce modèle implique plusieurs sous-catégories : les croyances, les opinions, les sentiments, les évaluations et les jugements. Enfin, les deux derniers modèles fréquemment cités sont le modèle OCC (Ortony et al., 1990) et le modèle *Appraisal in English* de Martin and White (2005). Le premier est à l'origine de la catégorisation proposée dans WordNet-Affect (Valitutti, 2004) et des règles formelles développées dans (Shaikh et al., 2009). Le second est utilisé dans les travaux de Neviarouskaya et al. (2010b), Bloom et al. (2007) et Whitelaw et al. (2005). Ces modèles, bien qu'issus de champs disciplinaires distincts, se définissent comme des théories de l'*appraisal* : ils mettent en valeur la notion d'acte évaluatif effectué par un agent envers une entité, un fait ou un événement. Cet intérêt porté au processus évaluatif offre l'avantage, pour une utilisation dans des travaux de SA, de rappeler qu'une expression de sentiment n'est pas définie uniquement par sa valence mais également par sa source et sa cible.

### 2.1.2 Le Modèle OCC (Ortony et al., 1990)

Nommé des initiales de ses trois concepteurs (Ortony, Clore, et Collins), le modèle OCC a pour objectif d'élaborer une théorie cognitive des émotions mettant en relief ce qui les provoque et ce qui les caractérise. A l'origine de ce modèle, est posée l'hypothèse que les individus prennent en compte trois aspects majeurs, lorsqu'ils perçoivent le monde : les événements, les agents et les objets. Les émotions sont définies comme des réactions à ces trois aspects du monde. Le modèle se structure ainsi en trois branches (Figure 2.2). La première branche inclut l'ensemble des émotions définies comme des réactions affectives aux événements évalués comme souhaitables ou non-souhaitables. La deuxième branche regroupe les émotions définies comme des réactions d'approbation ou de désapprobation face aux actions d'agents évalués selon leur mérite. Enfin, la troisième classe inclut des émotions relatives à une affection ou aversion pour des objets. Celles-ci sont causées par des réactions à un objet relativement à son attractivité.

La notion d'évaluation (*appraisal*) est centrale dans le modèle OCC. Les auteurs définissent trois variables sur lesquelles se fonde une évaluation : la désidérabilité, le mérite et l'attractivité. La désidérabilité se rattache à des émotions relatives à des événements et est évaluée par la prise en compte d'une structure complexe d'objectifs. Selon sa capacité à faciliter ou à pénaliser la réalisation de ces objectifs, l'événement sera perçu comme plus ou moins désirable. De même, le mérite qui se rattache à des émotions relatives aux actions d'agents, est définie selon une hiérarchie de standards. Les standards constituent l'ensemble des normes et des conventions servant de référence à un individu lorsqu'il formule une approbation ou une désapprobation. La plupart de ces standards sont acquis socialement. Enfin, l'attractivité, impliquée dans la manifestation d'émotions relatives à un objet, est évaluée relativement aux attitudes d'une personne.

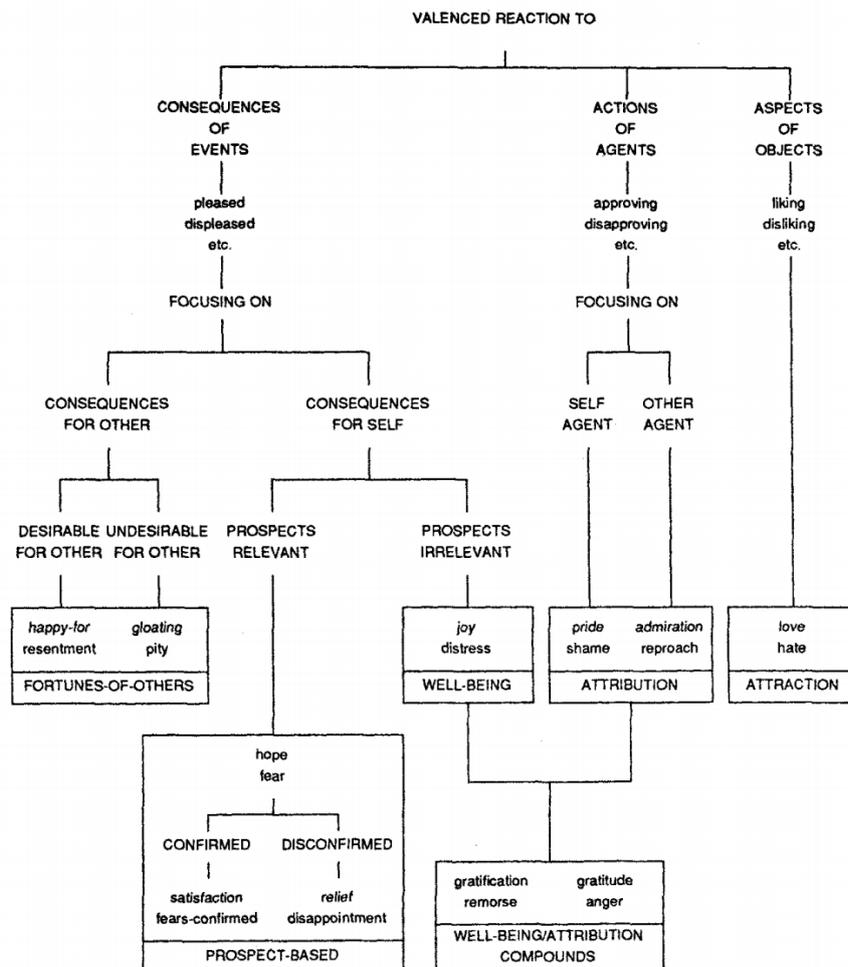


FIGURE 2.1 – Structure cognitive des émotions selon le modèle OCC

Elle détermine des émotions qui sont des réactions momentanées de goût ou de dégoût, et qui sont définies par les auteurs comme celles étant les plus immédiates, les plus spontanées et les moins affectées par des processus cognitifs.

### 2.1.3 *Appraisal in English* (Martin and White, 2005)

Cités dans plusieurs travaux d'analyse de sentiments, le modèle de (Martin and White, 2005) ne se contente pas de décrire des classes organisant les expressions de sentiment et d'opinion, il fournit également une analyse linguistique et recense un certain nombre d'indices lexicaux et de structures grammaticales. Martin & White se réclament de la *Linguistique Systémique Fonctionnelle* telle que théorisée par M.A.K Halliday dans (Halliday, 1973). Au cœur du modèle, se trouve cette volonté d'explicitier la manière dont les locuteurs construisent des jugements de valeur, attribuent leurs dires à des sources externes et modélisent leurs messages. Leur description du phénomène de l'évaluation tel qu'il s'exprime en anglais se base sur un modèle hiérarchique combinant trois sous-systèmes : celui de l'attitude, celui de la gradation et celui de l'engagement. A chacun de ces sous-systèmes sont associées des ressources lexico-grammaticales spécifiques.

La classe des attitudes « concerne nos sentiments, nos réponses émotionnelles, nos jugements concernant des comportements, et nos évaluations des choses [...] elle se réfère aux valeurs que les locuteurs communiquent à travers les jugements et les réponses émotionnelles/affectives qu'ils associent aux participants et aux processus » (Zhang, 2012a). Définies comme des pro-

cessus évaluatifs, les attitudes incluent "les signifiés par lesquels les textes/locuteurs attribuent une valeur ou une évaluation intersubjective aux participants et aux processus, que cela soit en rapport avec des réponses émotionnelles ou avec des systèmes de valeurs culturellement déterminés". D'un point de vue verbal, les attitudes sont des expressions référant à des processus évaluatifs à l'origine de ressentis émotionnels ou de références à des normes individuelles ou collectives. Trois sous-classes d'attitudes sont définies :

1. les *Affects* : cette catégorie regroupe des émotions ou des réactions émotionnelles. Les affects peuvent s'exprimer selon trois paradigmes : celui du bonheur vs. malheur, celui de la sécurité vs insécurité et enfin celui de la satisfaction vs insatisfaction.
2. les *Jugements* : cette sous-catégorie regroupe les évaluations portant sur des comportements ou des personnes. Elles sont formulées sur la base de principes normatifs et éthiques. Les jugements peuvent exprimer une sanction sociale. Ils réfèrent alors à des normes sociales codifiées (lois, codes ou règlements) qui édictent une manière de se comporter. Ce sont alors le respect de la décence et celui de la véracité qui seront en jeu. Mais les jugements peuvent également exprimer une forme d'estime sociale. Ils réfèrent alors à des attentes sociales ou à des systèmes de valeurs moins définis et évaluent leur cible selon son respect de la normalité, selon ses capacités et sa ténacité. Dans la mesure où les systèmes de valeur sont ancrés socialement, les jugements qui se réfèrent à eux doivent être considérés comme propres à une culture et une époque donnée.
3. les *Appréciations* : il s'agit d'évaluations formulées à l'égard d'artefacts, d'événements ou de processus. Les appréciations peuvent être des réactions aux choses (attirent-elles notre attention? Nous plaisent-elles?), à leur composition (sont-elles équilibrées? Quel est leur niveau de complexité?), ou à leur valeur (Sont-elles innovantes, authentiques, pertinentes, etc.)

Le modèle de Martin et White prend en charge la possible variation d'intensité du caractère positif ou négatif d'une attitude. A chaque attitude, est attribuée une propriété *gradation* qui s'articule autour de deux éléments majeurs :

1. la *Force* concerne la « modulation par les locuteurs de l'impact interpersonnel de leur message, en augmentent ou diminuent la force ou le volume. La force est une gradation opérant pour des catégories possédant une échelle de valeurs. Cette échelle de valeurs est exprimée de manière intrinsèque – différents degrés d'un même noyau de sens (ex : intelligent, et même génial), ou extrinsèque – différentes marques d'intensité ou de quantification (très, assez, trop, peu, etc.) » (Zhang, 2012a). La force peut connaître deux modes de réalisation : l'intensification et la quantification. L'intensification est applicable aux procès et aux qualités tandis que la quantification réfère au nombre, à la masse, à la proximité (temporelle ou spatiale), la distribution/étendue (temporelle ou spatiale).
2. le *Focus* désigne une gradation de l'appartenance à une catégorie sémantique. Le processus est celui d'un resserrement ou d'une ouverture des termes par lesquels l'appartenance à une catégorie est déterminée, comme par exemple : « un véritable ami/ presque un ami/ une espèce d'ami ».

Enfin dernier aspect, l'engagement : « l'engagement réfère aux ressources dont dispose le locuteur pour positionner sa voix par rapport aux divers énoncés et courants de pensées qui circulent [ainsi qu'aux] signifiés par lesquels les locuteurs reconnaissent ou ignorent la diversité des points de vue que leurs messages mettent en jeu et à travers lesquels ils négocient un espace interpersonnel pour leur(s) propre(s) position(s) » (Zhang, 2012a). Le locuteur a la possibilité de reconnaître ou non l'existence de positions autres que la sienne en faisant usage soit de *monoglossie*, soit d'*hétéroglossie*. La monoglossie implique la mise sous silence des autres positions dialogiques, comme par exemple la phrase (i) ci-dessous dans laquelle rien n'indique que la position prise est celle propre au locuteur. L'hétéroglossie inscrit au cœur même de l'énoncé

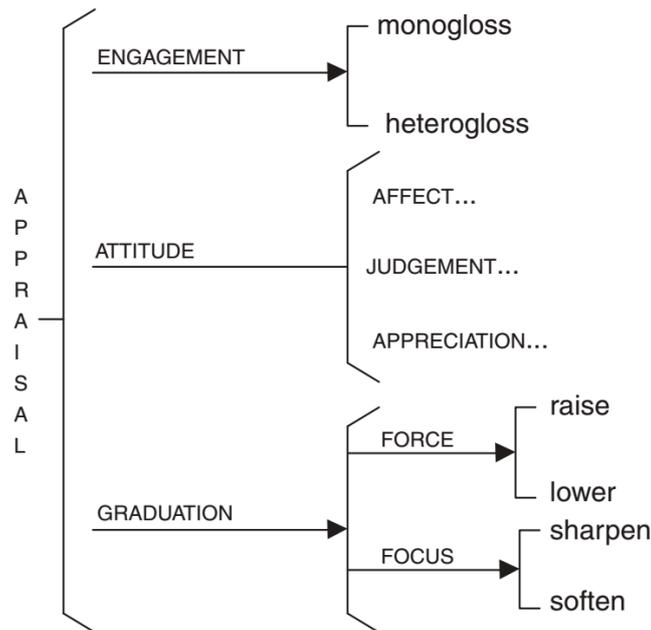


FIGURE 2.2 – Vue d'ensemble du modèle de Martin and White (2005)

l'existence d'autres voix. Dans la phrase (ii), « à mon avis » indique une prise de position propre au locuteur et la possible existence d'une altérité.

- (i) « Cette décision est mauvaise »
- (ii) « A mon avis, cette décision est mauvaise »

#### 2.1.4 Bilan

Si la définition de la notion de sentiment ainsi que l'emploi des termes y référant varient d'une étude à une autre, des modèles théoriques solides existent et permettent de problématiser l'approche. Plus fréquemment issus de la psychologie cognitive, ces modèles offrent une catégorisation et caractérisation fine du phénomène étudié. De plus, lorsqu'ils modélisent des sentiments et des émotions comme des processus évaluatifs (Ortony et al., 1990), ils font de l'expérienceur et du déclencheur des éléments constitutifs et indispensables à la compréhension du phénomène. Parmi ces différents modèles, le modèle de Martin and White (2005) se distingue par son approche linguistique. Les auteurs s'intéressent davantage aux expressions verbales qu'aux processus cognitifs auxquelles ces dernières réfèrent. Leur objectif est d'analyser les ressources verbales dont dispose le locuteur pour exprimer des attitudes. Cette spécificité nous a encouragée à choisir ce modèle comme fondement de notre approche. Notre objectif étant de nous concentrer exclusivement sur le contenu verbal, un modèle décrivant les ressources lexicogrammaticales nécessaires pour exprimer des attitudes nous est apparu comme pertinent. La nature des expressions d'attitude dans le cadre de conversations humain-agent sera étudiée en détail dans les chapitres 3 4 et 5.

## 2.2 Méthodes en analyse de sentiments

Cette section propose une description des différentes méthodes utilisées en analyse automatique de sentiments. Après un aperçu de l'évolution des objectifs et des méthodes, nous fournissons une description des travaux s'attachant à l'analyse détaillée des expressions de sentiment et à la production d'une description plus fine de leurs différentes variables (polarité, cible et source).

### 2.2.1 Aperçu des objectifs et méthodes en analyse de sentiments

Depuis ses débuts, le domaine de l'analyse de sentiments a connu différentes mutations, réorientant les objectifs et les méthodes proposées. Au cours de sa première décennie, les principaux travaux se sont attachés à la résolution de problèmes de classification : classification binaire en terme de polarité, positif et négatif, d'items lexicaux, de phrases et de textes. Concernant le lexique, le travail de Hatzivassiloglou and McKeown (1997) présente une approche innovante en proposant un modèle de régression non linéaire pour déterminer l'orientation sémantique (positive ou négative) d'adjectifs conjoints. La détermination automatique de l'orientation sémantique des mots est également au cœur du travail de Turney and Littman (2002) où celle-ci est calculée sur la base de la force de l'association du mot candidat avec sept mots positifs et sept mots négatifs. Les auteurs utilisent une méthode d'apprentissage non supervisé qui se fonde sur le calcul de l'information mutuelle (Church and Hanks, 1990) entre le mot candidat et les mots du lexique. Le succès des travaux portant sur la classification d'items lexicaux a ainsi encouragé la création de ressources lexicales comme SentiWordNet (Baccianella et al., 2010), EmotiWord (Malandrakis et al., 2011), le dictionnaire LIWC (Pennebaker et al., 2001), WordNet Affect (Valitutti, 2004), ou encore Sentiful (Neviarouskaya et al., 2011).

Au-delà du lexique, la classification de textes et de phrases est également un challenge important en SA. Pour attribuer une polarité globale à un texte ou à une phrase, différents algorithmes d'apprentissage automatique ont été utilisés (Pang et al., 2008), (Liu, 2012). Pour évaluer la pertinence des méthodes d'apprentissage supervisé, (Pang et al., 2002) offre une analyse comparative des performances de trois algorithmes d'apprentissage (*classifieur naïf bayésien*, *Machine à Vecteur de Support*, classifieur d'entropie maximale) pour la classification de revues de films. Dans le cadre de l'emploi de ces techniques, la recherche porte également sur la définition de caractéristiques ou d'indices (*features*) pertinents et efficaces (Liu, 2012) : fréquence des mots (unigramme ou n-gramme), catégories grammaticales, information sur les dépendances syntaxiques, lexique évaluatif.

Ces méthodes ont permis de résoudre différents problèmes, néanmoins la classification *globale* de textes ou de phrases avec des méthodes d'apprentissage supervisé montre certaines limites :

- L'attribution d'une polarité à l'ensemble du texte ou de la phrase concernant un objet spécifique ne doit pas mettre de côté l'analyse des opinions portant sur les différents aspects de cet objet. Il est en effet possible qu'un utilisateur puisse évaluer positivement un film sans pour autant aimer tous ces différents aspects (scénario, mise en scène, jeu d'acteur, etc.). Dans la même perspective, ces méthodes ne permettent pas toujours une identification précise de la cible lorsque celle-ci n'est ni l'objet du texte, ni l'un de ses aspects.
- Les méthodes supervisées demandent de larges corpus annotés qui ne sont pas toujours disponibles et requièrent un grand investissement pour être créés. De plus, les modèles ne permettent pas toujours de contrôler et d'interpréter leurs résultats, car le phénomène est complexe et difficile à annoter.

Un certain nombre de travaux ont ainsi montré une volonté d'amélioration de l'analyse des différentes caractéristiques des expressions de sentiments. Dans la section 2.2.2, nous présentons un ensemble de travaux s'intéressant au principe de compositionnalité pour produire des

méthodes plus efficaces pour le calcul de la polarité. Quittant l’unité du texte ou de la phrase, ils s’attachent davantage à celle de l’expression de sentiment et aux différents mécanismes sémantiques qui s’y jouent. Cet intérêt porté à l’expression même se retrouve également dans les travaux se réclamant de l’*Aspect-Based Sentiment Analysis* (décrits dans Section 2.2.3) qui, afin de pouvoir identifier les sentiments exprimés au sujet des différents aspects d’une cible, proposent des méthodes permettant une analyse plus fine du contenu textuel.

## 2.2.2 Compositionnalité et polarité

Afin de mieux prendre en charge la polarité et la gradation des expressions de sentiment (Dzogang et al., 2010), de nombreuses méthodes se sont appuyées sur le principe de compositionnalité. Dans son acception la plus commune, ce principe repose sur l’hypothèse que le sens d’une expression complexe est défini par celui des éléments qui la composent et par les règles qui organisent leurs combinaisons (Szabó, 2012). Deux types de méthode peuvent être identifiés : (i) des méthodes proposant des règles expert et cherchant à définir de manière exhaustive des patrons syntaxiques auxquels sont associées des règles de calcul de la polarité ; (ii) des méthodes hybrides associant des algorithmes d’apprentissage automatique à des représentations symboliques de la phrase.

**Développement de règles expert** Les méthodes reposant entièrement sur le développement de règles expert se concentrent sur une description formalisée des phénomènes linguistiques qu’elles cherchent à détecter. En analyse de sentiments, l’objectif est donc de définir un nombre important de patrons syntaxiques et de règles sémantiques utiles à la détection de sentiments et d’opinion. Dans Shaikh et al. (2007), les auteurs proposent une méthode de détection des sentiments s’organisant en trois temps. Tout d’abord, une phase d’analyse sémantique, s’intéressant aux relations de dépendance de chaque phrase et produisant pour chacune d’elles un ensemble de triplets (sujet, verbe, objet). Sur la base de cette analyse sémantique et actancielle, une polarité est assignée à chaque triplet. Enfin, dans un troisième temps, une polarité globale pourra être assignée à chaque phrase analysée. Pour calculer la polarité d’un triplet, le système utilise un ensemble de lexiques définissant la polarité de chaque mot hors-contexte. Sur la base de cette information et de l’analyse sémantique effectuée pendant la première phrase, le système définit avant tout une polarité pour chaque membre du triplet avant de définir celle du triplet lui-même.

$ADJ_{pos} + (CON_{neg} \text{ or } NE_{neg}) \rightarrow \text{NEG}$ (ex : strong cyclone) $ADJ_{pos} + (CON_{pos} \text{ or } NE_{pos}) \rightarrow \text{POS}$ (ex : brand new car) $ADJ_{neg} + (CON_{pos} \text{ or } NE_{pos}) \rightarrow \text{NEG}$ (ex : ugly princess) $ADJ_{neg} + (CON_{pos} \text{ or } NE_{pos}) \rightarrow \text{NEG}$ (ex : ugly witch)
--

TABLEAU 2.1 – Règles pour l’attribution d’une polarité aux groupes nominaux (Shaikh et al., 2007)

Le tableau 2.1 montre des exemples de règles appliquées au niveau du groupe nominal. Selon la polarité de l’adjectif, positive ( $ADJ_{pos}$ ) ou négative ( $ADJ_{neg}$ ) et de l’entité nommée (notée  $NE_{pos|neg}$  selon la polarité véhiculée) ou du concept ( $CON_{pos|neg}$ ) qu’il qualifie, il est alors possible de définir la polarité finale du groupe nominal (notée  $\rightarrow POS|NEG$ ). Une fois définie la polarité de chaque membre du triplet, le système définit une polarité pour le couple <action, objet>. La polarité du triplet est définie au regard de la polarité attribuée à son élément *sujet* et au couple <action, objet> (Tableau 2.2).

$(CON_{pos} \text{ or } NE_{pos}) + \langle \text{Action} - \text{Object} \rangle_{pos} \rightarrow \text{Triplet}_{pos}$ (e.g., the professor explained the idea to his students.)
$(CON_{pos} \text{ or } NE_{pos}) + \langle \text{Action} - \text{Object} \rangle_{neg} \rightarrow \text{Triplet}_{neg}$ (e.g., John rarely attends the morning lectures.)
$(CON_{neg} \text{ or } NE_{neg}) + \langle \text{Action} - \text{Object} \rangle_{pos} \rightarrow \text{Triplet}_{neg}$ (e.g., the robber appeared in the broad day light.) to process further.
$(CON_{neg} \text{ or } NE_{neg}) + \langle \text{Action} - \text{Object} \rangle_{neg} \rightarrow \text{Triplet}_{neg}$ A verifier (e.g., the strong cyclone toppled the whole city.)

TABLEAU 2.2 – Règles pour l’attribution d’une polarité aux triplets (Shaikh et al., 2007)

Moilanen and Pulman (2007) proposent un modèle intégrant le principe de compositionnalité pour le calcul de la polarité. Il prend en entrée deux constituants syntaxiques dont il calcule simultanément la polarité ainsi que la polarité du constituant d’ordre supérieur issu de leur combinaison. Le calcul de la polarité des constituants d’ordre supérieur se fonde sur des calculs sémantiques génériques. Trois cas de figure sont envisagés :

1. **Propagation de la charge affective** : la polarité d’un constituant non-neutre est transmise à un constituant neutre et au constituant d’ordre supérieur.
2. **Inversion de polarité** : un constituant inverse la polarité d’un constituant non-neutre.
3. **Conflit de polarité et résolution de conflit** : les deux membres d’un segment ont deux polarités différentes. Dans ce cas, une règle sémantique est appliquée pour résoudre ce conflit. Cette résolution est établie sur la base d’un statut de subordonné ou subordonnant attribué à chaque constituant.

Les auteurs utilisent un lexique dans lequel à chaque entrée lexicale est attribué l’un des labels suivants : *default*, *positive*, *negative*, *reverse*. Ces labels sont également ceux qui seront attribués à chaque constituant de la phrase au cours de l’analyse. Afin de pouvoir correctement calculer la polarité de chaque segment, les auteurs se fondent sur la nature de leur constructions grammaticales.

Dans (Klenner et al., 2009), les auteurs décrivent une méthode symbolique fondée sur l’emploi des règles sémantiques et de patrons syntaxiques pour la classification de phrases subjectives. Trois catégories sont utilisées pour les différents niveaux d’analyse : positif (POS), négatif (NEG) et neutre (NEU). Pour justifier leur méthode d’analyse et de classification, les auteurs affirment que les règles de compositionnalité déterminant la polarité d’une phrase sont assez peu nombreuses pour pouvoir être implémentées facilement. Pour faciliter l’identification d’indices lexicaux, ils ont constitué manuellement un lexique de 8000 mots. Le lexique est construit à partir de celui initialement utilisé dans (Wilson et al., 2005). Les règles développées pour le calcul de la polarité s’appliquent aux différents niveaux de la structure hiérarchique de la phrase. Elles sont conçues comme des règles de réécriture prenant en entrée des constituants de niveau inférieur et produisent en sortie la polarité du constituant de niveau supérieur. Le tableau 2.3 décrit les règles appliquées au niveau du groupe nominal restreint tandis que le tableau 2.4 présentent celles appliquées au niveau du groupe nominal étendu, i.e. avec complément du nom.

ADJ	NOUN	→	NP	Exemples
NEG	POS	→	NEG	<i>disappointed hope</i>
NEG	NEG	→	NEG	<i>a horrible liar</i>
POS	POS	→	POS	<i>a good friend</i>
POS	NEG	→	NEG	<i>a perfect misery</i>
POS	NEU	→	POS	<i>a perfect meal</i>
NEG	POS	→	NEG	<i>a horrible meal</i>

TABLEAU 2.3 – Calcul de la polarité au niveau du groupe nominal restreint (Klenner et al., 2009)

NP	Prep	NP	→	PP	Exemples
POS	<i>to</i>	NEG	→	POS	<i>solution to my problem</i>
POS	<i>for</i>	POS	→	POS	<i>hope for relief</i>
NEG	<i>of</i>	NEG	→	NEG	<i>pain of disappointment</i>
NEG	<i>of</i>	POS	→	NEG	<i>lost of hope</i>

TABLEAU 2.4 – Calcul de la polarité au niveau du groupe nominal étendu (Klenner et al., 2009)

Neviarouskaya et al. (2010b) présentent un système de classification des phrases – @AM (Attitude Analysis Model). Ce système s’appuie sur le modèle des émotions de Izard (Izard, 1977) et sur le modèle *Appraisal* de Martin and White (2005). Neuf catégories d’affect sont ainsi utilisées : *anger, disgust, fear, guilt, interest, joy, sadness, shame, surprise*. Du côté des appréciations et des jugements, le système réfère à quatre labels combinant polarité et type d’attitude : jugement positif et négatif (*NEG Jug, POS Jug*), appréciation positive et négative (*NEG App* et *POS app*). Le système est également capable d’assigner une intensité à chaque phrase traitée. Reposant sur une approche compositionnelle, le système définit la polarité globale de la phrase sur la base de la polarité de chacun de ces constituants et de règles combinatoires. Afin de disposer d’une polarité pour chaque mots, les auteurs utilisent un lexique (*Affect Database*). Pour calculer la polarité d’une phrase (voir le tableau 2.5), le système développé fait jouer, à différents niveaux de l’analyse, six règles : *inversion, aggrégation, propagation, dominance, neutralisation*, et enfin, *intensification*.

**Méthodes hybrides** Depuis quelques années, un certain nombre de méthodes dites *hybrides* ont été proposées pour améliorer les performances des systèmes de détection. Ces méthodes combinent des algorithmes d’apprentissage automatique et des règles expert incluant des patrons syntaxiques ou des informations sur les relations de dépendance structurant la phrase. Leur intérêt est de tirer parti des deux types d’approches présentées ci-dessus : la puissance de calcul et la généralité des algorithmes d’apprentissage automatique ainsi que la précision d’analyse des approches par règles expert.

Yang and Cardie (2013) proposent une méthode qui se fonde sur l’identification conjointe des différents éléments composant une opinion – l’expression d’opinion, la source (*holder*) et la cible (*target*) – et des relations posées entre elles – relation *is-about* associant l’expression et la

Règles	Modificateurs/Condition d'Application	Exemples
Inversion	1) Mot de négation inverse la polarité du mot porteur d'attitude 2) Adverbe de doute inverse la polarité du mot porteur d'attitude 3) Mot fonctionnel inverse la polarité du mot porteur d'attitude	<i>never</i> + POS( <i>succeed</i> ) → NEG( <i>never succeed</i> ) <i>scarcely</i> + POS( <i>relax</i> ) → NEG( <i>scarcely relax</i> ) <i>reduced</i> + POS( <i>enthusiasm</i> ) → NEG( <i>reduced enthusiasm</i> )
Aggrégation	Mots avec polarités opposées	POS( <i>beautiful</i> ) + NEG( <i>fight</i> ) → POS( <i>beautiful fight</i> )
Propagation	Propagation de la polarité du verbe vers les arguments	POS( <i>to admire</i> ) + NEUTR( <i>his behaviour</i> ) → POS( <i>his behaviour</i> )
Dominance	Polarité du verbe domine celle de son objet	NEG( <i>to deceive</i> ) + POS( <i>hopes</i> ) → NEG( <i>to deceive hopes</i> )
Neutralisation	Modificateur ou opérateur de condition neutralise la polarité du mot porteur d'attitude	<i>despite</i> + NEG( <i>worries</i> ) → NEUTR( <i>despite worries</i> )
Intensification	1) Adverbe de degré accentue la polarité du mot porteur d'attitude 2) Comparatif accentue la polarité du mot porteur d'attitude	POS( <i>happy</i> ) vs. POS++ ( <i>very happy</i> ) NEG( <i>sad</i> ) vs. NEG++ ( <i>saddest</i> )

TABLEAU 2.5 – Règles pour le calcul de la polarité (Neviarouskaya et al., 2010a)

cible et la relation *is-from* associant la source à l'opinion exprimée. L'identification des entités constitutives d'une opinion est envisagée comme un problème de *sequence tagging* et l'extraction des relations comme une tâche de classification binaire. S'intéressant à la manière dont se structure une expression d'opinion, les auteurs utilisent, comme données d'entraînement, un corpus où les différents composants des opinions ainsi que leurs relations sémantiques ont été annotés manuellement. Pour identifier les composants d'une opinion, ils utilisent des champs conditionnels aléatoires (*Conditional Random Fields*). Les relations entre ces derniers sont ensuite établies sur la base de patrons syntaxiques : "we selected the most common patterns of the shortest dependency paths between an opinion candidate and an argument candidate in our dataset, and include all pairs of candidates that satisfy at least one dependency pattern".

Afin de mieux gérer les effets de la compositionnalité sémantique, Socher et al. (2013) proposent une méthode d'analyse ascendante pour la classification des phrases. Cette méthode se fonde sur un modèle de *Recursive Neural Tensor Networks* qui calcule une représentation vectorielle pour des syntagmes de longueur variable et de types syntaxiques différents. Ces représentations sont ensuite utilisées pour classer chaque syntagme. Quand un *n-gram* est donné au modèle, il est analysé comme un arbre binaire et chaque feuille correspondant à un mot est représentée par un vecteur. Le système calcule ensuite les vecteurs parents en utilisant un processus d'analyse bottom-up et différentes fonctions de calcul de compositionnalité. Les vecteurs parents sont ensuite donnés comme *features* au classifieur.

(Rentoumi et al., 2010) proposent un système de détection des sentiments se concentrant sur les expressions figuratives. Les auteurs définissent ce type d'expressions comme reposant sur un processus sémantique d'extension métaphorique, de restriction de l'usage du mot. Les règles linguistiques sont ajoutées pour pouvoir gérer, selon la terminologie des auteurs, des cas marginaux d'expressions figuratives où la compositionnalité sémantique doit être prise en compte pour établir correctement la polarité de l'expression. La partie apprentissage est quant à elle utilisée pour la gestion des cas non-marginaux. L'analyse est effectuée en quatre temps : (i)

une désambiguïstation sémantique au niveau du mot ; (ii) l’assignation d’une polarité à chaque mot désambiguïté, approche *machine learning* exploitant des graphes fondés sur des *character n-grams* (HMMs). Ils ont deux fonctions : la première, définir le seuil permettant de diviser le phrases en marginales/non-marginales ; la seconde, d’attribuer une polarité (positive/négative) aux phrases non-marginales. Au niveau phrastique, la polarité est attribuée par les HMMs aux phrases non-marginales, et par le système à base de règles linguistiques aux les phrases figuratives.

### 2.2.3 Approches *aspect-based* : vers une meilleure analyse des cibles

Comme l’explique Liu (2012), la classification de textes ou de phrases véhiculant une opinion est souvent insuffisante pour certaines applications puisque les cibles ou l’assignation de sentiments à ces mêmes cibles ne sont pas identifiées. Même en formulant l’hypothèse qu’un même texte évalue une seule entité, l’opinion positive généralement exprimée ne signifie pas nécessairement que l’auteur a une opinion positive de toutes les facettes (*aspects*) de cette entité. Ainsi, les travaux se réclamant de l’approche *aspect-based* ont pour objectif d’identifier les différentes facettes d’une cible et de déterminer la nature des sentiments envers chacune d’elles. Comme l’explique He et al. (2017), deux types de sous-tâches peuvent se rattacher au domaine de l’extraction d’aspects : d’un côté l’extraction des termes référant à un aspect particulier, de l’autre le regroupement (*clustering*) de termes synonymes au sein de catégories caractérisant chaque aspect.

Les premiers travaux dédiés à l’extraction d’aspects reposent pour la plupart sur des règles experts ((Hu and Liu, 2004), (Zhuang et al., 2006), (Su et al., 2009)). Parmi les travaux plus récents, nous pouvons citer (Qiu et al., 2011). Afin de détecter conjointement la cible et l’expression d’opinion qui l’évalue, (Qiu et al., 2011) se concentrent sur les relations syntaxiques pouvant exister entre ces deux éléments. Ces relations sont formalisées par des grammaires de dépendances. Plus récemment, les travaux du domaine ont montré un intérêt grandissant pour les méthodes par apprentissage non-supervisé ((He et al., 2017), (Brody and Elhadad, 2010)). Certain d’entre eux utilisent même des ressources linguistiques. Ainsi,(Wang et al., 2016) élaborent un modèle de réseau de neurones sur la base de l’arbre de dépendance de chaque phrase (Socher et al., 2010). Chaque mot est appréhendé dans son contexte phrastique. De la sorte, les relations de dépendance existant entre les termes véhiculant une opinion et ceux référant à la cible peuvent être apprises par l’algorithme. Le résultat est ensuite exploité par un modèle de CRFs pour l’apprentissage. Leur modèle est enfin évalué sur le corpus SemEval 2014 (Pontiki et al., 2016).

### 2.2.4 Bilan

Dans cette section, nous avons présenté différents travaux qui diffèrent tant par leurs objectifs que par les méthodes utilisées. Concernant les objectifs, tandis que certains travaux se concentrent sur des classifications de données massives issues du web, d’autres cherchent à extraire et à individualiser les expressions de sentiment et se concentrent sur l’analyse de caractéristiques précises : la polarité et la cible notamment. Concernant les méthodes, trois types émergent : des méthodes intégrant des algorithmes d’apprentissage supervisé, des méthodes intégrant des règles expert et des méthodes hybrides, combinant exploitation de règles linguistiques et apprentissage automatique. Chacune d’elles présentent un certain nombre d’avantages – précision et robustesse, pour les méthodes par apprentissage et finesse d’analyse pour les méthodes à base de règles – et semblent particulièrement adaptées à un certain type d’usage (classification vs analyse sémantique et extraction d’aspects).

## 2.3 Positionnement : analyse de sentiments pour les interactions humain-agent conversationnel animé

Dans le domaine de l'interaction humain-machine, seuls quelques travaux ont pu s'intéresser à l'analyse du contenu verbal pour aider à la détection d'états cognitifs (Zhang et al., 2006) ou émotionnels ((Litman and Forbes-Riley, 2004), (Yildirim et al., 2011), (Smith et al., 2011b)). Pour la plupart, il s'agit essentiellement d'extraire des indices lexicaux en support à l'analyse d'indices non-verbaux. Par exemple, afin de détecter automatiquement les attitudes de frustration et de politesse chez des enfants interagissant avec des personnages de jeux vidéo, Yildirim et al. (2011) utilisent et combinent différents indices : des caractéristiques acoustiques, des indices lexicaux et des indices contextuels. Pour l'analyse et la modélisation, les auteurs ont utilisé la base de donnée *Children's interactive multimedia Project* (Narayanan and Potamianos, 2002). Pour extraire les marqueurs lexicaux qui aideront à la détection de l'état émotionnel, les auteurs utilisent deux différentes approches. La première associe un *algorithme de classification bayésien* et une mesure de *l'information mutuelle*. La seconde repose sur l'hypothèse que les phrases de l'utilisateur exprimant les mêmes émotions partagent le même espace sémantique. Pour calculer la distance émotionnelle entre chaque phrase, les auteurs utilisent une analyse sémantique latente. Les résultats de l'expérimentation montrent que les indices lexicaux ont un pouvoir de discrimination plus important que les traits acoustiques et contextuels pour la détection de la politesse, tandis que les traits contextes et acoustiques sont plus performants pour la détection de la frustration. De plus, la fusion des indices lexicaux, acoustiques et contextuels apportent une amélioration significative des résultats.

La seule étude utilisant un module analysant des segments plus complexes que les items lexicaux (Smith et al., 2011b) n'adapte pas celui-ci au discours conversationnel. Pour le développement de notre système, nous avons décidé de nous consacrer uniquement au contenu verbal et de nous attacher à construire un modèle prenant en compte le contexte conversationnel. Si à long terme, il sera possible et nécessaire d'analyser les expressions verbales de sentiments conjointement aux indices multimodaux (Soleymani et al., 2017), nous considérons qu'il est avant tout nécessaire de produire une modélisation plus complète de ces expressions. Le travail présenté ici n'a pas pour but premier l'aide à la détection des émotions de l'utilisateur, mais bien à comprendre et modéliser la manière dont ce dernier utilise le contenu verbal pour décrire des ressentis, des goûts, des préférences.

### 2.3.1 Choix du modèle théorique

Cette attention portée au contenu verbal est déterminante pour le choix du modèle théorique de référence. Le modèle doit fournir une description détaillée de expressions de sentiment ainsi qu'une catégorisation fine. Une simple distinction posée entre des énoncés positifs et des énoncés négatifs ne serait donc suffire. Pour cette raison, notre choix s'est arrêté sur le modèle de Martin and White (2005) décrit Session 2.1.3. Il nous a en effet semblé pertinent de choisir un modèle linguistique et non un modèle issue de la psychologie. Les travaux de psychologie cognitive offrent une description détaillée des affects, émotions et sentiments. Néanmoins, leur objectif est bien d'appréhender ces phénomènes en tant que mécanismes psychologiques, ils apparaissent donc comme lacunaires en ce qui concerne la description de phénomènes verbaux. Le modèle de Martin and White (2005), au contraire, a pour objectif de décrire verbalement les expressions de sentiment et fournit une analyse approfondie de leur lexique ainsi que des structures phrastiques et dialogiques au sein desquelles ils s'intègrent. En outre, ce modèle propose une catégorisation fine de ces phénomènes, ce qui est essentiel. En posant une distinction entre des expressions référant à des affects et d'autres référant à des jugements de valeur, ce modèle offre la possibilité de sélectionner parmi le vaste ensemble des expressions de sentiments, la catégorie qui nous semblera pertinente dans le cadre d'interactions spécifiques avec l'agent.

### 2.3.2 Choix de la méthode

L’utilisation d’un modèle linguistique permettra également de pouvoir créer des ressources linguistiques (lexiques, règles) en s’appuyant sur une solide base théorique. Étant donné le caractère inédit de l’analyse verbale des sentiments dans les interactions humain-agent, nous considérons que le développement des règles expert permettra de mieux appréhender et de formaliser le phénomène dans ce type de contexte. De plus, en l’absence de corpus conversationnels comprenant des annotations fines des expressions de sentiments, l’exploitation de méthodes par apprentissage supervisé nous paraît exclue. De plus, si les méthodes de classification des textes et des phrases présentées Section 2.2 pourraient a minima être utilisées pour attribuer une polarité globale au tour de parole, cette solution n’est pour nos besoins pas suffisante. Il est nécessaire que le système développé puisse individualiser les expressions de sentiment afin de permettre aux système agent de pouvoir y répondre. Pour cette raison, les trois attributs principaux des expressions de sentiment - source, cible et polarité - doivent être définis le plus précisément possible. Concernant la source, il est indispensable de pouvoir distinguer les sentiments dont l’utilisateur est la source de celles dont la source est une autre personne (personne ayant exprimé par exemple un jugement de valeur et dont l’utilisateur rapporterait les propos). Pour la polarité, il nous paraît nécessaire de pouvoir distinguer, à l’intérieur même du tour de parole, les différentes expressions ainsi que leur polarité si celle-ci diffère, or les méthodes prenant en compte la compositionnalité ont montré leur efficacité pour le calcul de la polarité. Si à long-terme, ces règles pourront être exploitées par un système hybride, leurs performances doivent être dans un premier temps évaluées. C’est en partie l’objectif de nous fixons dans notre thèse.

## **Deuxième partie**

# **Fondements linguistiques : les expressions d'attitude dans les conversations**



---

## Synthèse de la partie II

Cette partie a pour objectif d'analyser les expressions d'attitude telles qu'elles se manifestent dans le contexte de conversations humain-agent et d'asseoir les fondements linguistiques nécessaires au développement du système présenté dans la troisième partie du mémoire. Trois niveaux d'analyse sont envisagés : le tour de parole, la paire adjacente et la séquence thématique. Chacun de trois chapitres composant cette partie se concentre sur l'un de ces cadres. Chacun d'eux suit la même approche. Dans un premier temps, nous posons les bases théoriques issues de la littérature permettant de mieux appréhender le phénomène et/ou le cadre d'analyse. Dans un deuxième temps, nous proposons une analyse des mécanismes syntaxiques, sémantiques et/ou pragmatiques spécifiques aux expressions d'attitude dans le contexte considéré. Dans un troisième, nous proposons une formalisation de ces mécanismes servant de base aux règles et aux grammaires développées dans la troisième partie du mémoire. Chacune des analyses proposées dans cette partie se fonde sur des observations menées sur nos deux corpus d'étude : le corpus Semaine (McKeown et al., 2011) et le corpus de négociation (Gratch et al., 2016).

Le chapitre 3 a pour cadre d'analyse le tour de parole de l'utilisateur. Nous nous intéressons aux expressions simples d'attitude (ESA), nommées de la sorte par opposition aux expressions d'attitude strictement collaboratives (EASC) étudiées dans le chapitre 4. Contrairement aux EASC, les ESA ne nécessitent que le tour de parole pour être reconnues comme expressions d'attitude. Dans un premier temps, le chapitre se concentre sur l'analyse des patrons syntaxiques des attitudes et des règles sémantiques trouvés dans la littérature. Après avoir expliqué leur fonctionnement, nous proposons une extension afin de pouvoir cumuler informations syntaxiques et sémantiques relatives notamment au calcul de la polarité et de la définition des rôles sémantiques de *source* et de *cible*.

Le chapitre 4 a pour objectif d'interroger le caractère participatif des expressions d'attitude en exploitant la notion de contribution dialogique telle qu'elle est définie par (Clark and Schaefer, 1989). Dans le cadre de la paire adjacente, nous analysons la manière dont les attitudes exprimées par l'utilisateur fonctionnent avec le tour de parole de l'agent qui les précède. Nous nous attachons particulièrement à la description des EASC. Ces expressions ont la particularité de ne pouvoir être détectées que par la prise en compte conjointe du tour de parole de l'agent et de celui de l'utilisateur. Typiquement, elles reposent sur la formulation d'une question ou d'une assertion véhiculant une attitude (*Do you like outdoors activities?*) et d'une réponse permettant de valider ou d'invalider l'attitude (*Yes, I do*). Après une analyse détaillée des EASC, nous formalisons les mécanismes sémantiques et pragmatiques, existant entre le tour de parole de l'agent et celui de l'utilisateur, permettant de définir les valeurs des attributs *cible*, *source* et *polarité*.

Le chapitre 5 a pour cadre d'analyse la séquence thématique. Nous y analysons la structure thématique des conversations humain-agent ainsi que les relations sémantiques pouvant exister entre la cible et le thème courant. L'objectif de cette analyse est de fournir une méthodologie permettant d'asseoir la détection des attitudes sur les informations sur le thème fournies par le moteur de dialogue.



## Chapitre 3

# Les attitudes à l'échelle du tour de parole de l'utilisateur

### Sommaire

---

3.1	Description syntaxique et sémantique des attitudes dans la littérature . . . . .	<b>31</b>
3.1.1	Patrons syntaxiques . . . . .	31
3.1.2	Règles sémantiques . . . . .	31
3.2	Formes syntaxiques, rôles sémantiques et calcul de la polarité . . . . .	<b>34</b>
3.2.1	Structures syntaxiques et identification de la source et de la cible . . .	34
3.2.2	Structures syntaxiques et règles de polarité . . . . .	38
3.3	Spécificités liées au contexte conversationnel . . . . .	<b>41</b>
3.4	Conclusion du chapitre 3 . . . . .	<b>43</b>

---

C E chapitre est consacré à l'étude du tour de parole. Notre volonté est de définir les structures logico-sémantiques pouvant être exploitées par l'utilisateur pour exprimer une attitude à l'intérieur du tour de parole. Cette approche s'intéresse aux mécanismes sémantiques et syntaxiques opérant au sein des énoncés. Dans le chapitre suivant (chapitre 4), elle cèdera la place à une approche se concentrant davantage sur la manière dont les énoncés s'articulent entre eux.

L'objectif ici est de définir les patrons syntaxico-sémantiques des expressions d'attitudes. Notre approche rejoint celle de travaux antérieurs qui, pour formaliser les expressions de sentiments, s'attachent à leur forme syntaxique et sémantique. Cette caractérisation syntaxique et sémantique nous apparaît comme une étape nécessaire pour la création d'un ensemble de règles linguistiques de détection. Ce chapitre s'articule en trois temps. Dans un premier temps (section 3.1), nous présenterons les patrons syntaxiques et les règles sémantiques identifiés dans la littérature. Nous nous concentrerons particulièrement sur les patrons syntaxiques décrits par Martin and White (2005) et par Bednarek (2009) et sur les règles sémantiques proposées par Neviarouskaya et al. (2010b). Dans un deuxième temps (Section 3.2), nous proposerons une reformulation et une extension des patrons de Martin and White (2005) et de Bednarek (2009). Cette reformulation aura pour objectif de mettre en évidence la manière dont la valeur des attributs polarité, source et cible est définie par la structure logico-sémantique de la phrase, afin par la suite de simplifier l'écriture des grammaires utilisées dans le système. Les nouveaux patrons intègrent ainsi les rôles sémantiques de source et de cible, et indiquent le mode d'application des règles sémantiques de polarité aux différents niveaux de l'énoncé. Cette deuxième section sera également pour nous l'occasion de présenter des patrons n'apparaissant pas dans la typologie de Martin and White (2005). Enfin, dans un troisième temps, nous procéderons à une exemplification sur deux extraits des deux corpus de d'étude.

Publications associées à ce chapitre :

- Langlet, C. and Clavel, C. (2015b). Improving social relationships in face-to-face human-agent interactions: when the agent wants to know user's likes and dislikes. In *Proceedings of the 53rd Annual Meeting of the Association for Computational Linguistics and the 7th International Joint Conference on Natural Language Processing (Volume 1: Long Papers)*, volume 1, pages 1064–1073
- Langlet, C. and Clavel, C. (2014a). Modélisation des questions de l'agent pour l'analyse des affects, jugements et appréciations de l'utilisateur dans les interactions humain-agent. In *Proceedings of TALN 2014 (Volume 2: Short Papers)*, pages 479–484, Marseille, France. Association pour le Traitement Automatique des Langues

### 3.1 Description syntaxique et sémantique des attitudes dans la littérature

#### 3.1.1 Patrons syntaxiques

Martin and White (2005), et Bednarek (2009) dans leur continuité, recensent un certain nombre de patrons syntaxiques d'expressions d'attitude. Tous les patrons recensés accueillent une unité lexicale qui réfère au processus axiologique ou affectif en jeu dans l'expression d'une attitude. Cette lexie, ainsi que son contexte syntaxique, vont déterminer la sous-catégorie d'attitude à laquelle la phrase fait référence. Tous peuvent être décrits comme des variantes de la forme attributive canonique : Sujet + Vb attributif + Attribut du sujet. Dans ce type de structure, le verbe principal ne réfère pas à un procès. Il permet d'inscrire syntaxiquement la relation de qualification existant entre l'attribut (*wonderful* dans la phrase (i) ci-dessous) et le sujet (*This movie* dans la phrase (i)). Typiquement, c'est le verbe *to be* qui est utilisé comme verbe attributif. D'autres verbes peuvent néanmoins jouer ce rôle (*to seem, to get, to appear*).

(i) "This movie is wonderful!"

Selon leurs spécificités syntaxiques, nous pouvons répartir les patrons décrits par Martin and White (2005) et Bednarek (2009) en trois différents groupes (voir tableau 3.1) 1A, 1B et 2 qui se distinguent de la manière suivante :

- **Présence d'un verbe d'opinion** (principale différence entre les patrons du groupe 1 et ceux du groupe 2) : les verbes d'opinion ou de pensée, comme *to think, to consider, to see*, ne doivent pas être confondus avec les verbes référant à une attitude, comme *to like* ou *to hate*. Les premiers ne portent pas de polarité, ils servent simplement à indiquer qu'une opinion ou qu'un processus cognitif est à l'œuvre. Les seconds réfèrent à un processus d'évaluation axiologique ou affectif s'inscrivant sur un axe de polarité (positif vs. négatif).
- **La place de sujet est occupée soit par un nom** ("Horses are pretty to look at") **soit par un pronom indéfini** ("It was reasonable to incur that expense"). Cette caractéristique permet de distinguer les patrons du groupe 1A et ceux du groupe 1B. Pour le groupe 1B, le pronom a pour référent une proposition infinitive ("to hear her talking like that") ou participiale ("talking to you the other day") rattachées à l'attribut.

Bednarek définit un patron supplémentaire ne pouvant pas être intégré à l'un des groupes apparaissant dans le tableau 3.1. Selon l'auteur, ce patron, centré sur l'expérience du sujet, permet d'exprimer exclusivement des affects :

NP-Make-NP-(Feel)\*-AdjP

(ii) "This painting makes me feel happy"

#### 3.1.2 Règles sémantiques

Alors que l'analyse des expressions d'attitude menée par Bednarek ainsi que par Martin et White s'attache principalement aux formes de surface, le modèle présenté par Neviarouskaya et al. (2010b) s'intéresse davantage à leur fonctionnement sémantique. Cette étude définit un ensemble de règles sémantiques permettant d'interpréter les expressions d'attitude et de leur attribuer une polarité. La notion de compositionnalité est au cœur de ce modèle qui considère le sens d'une expression d'attitude comme défini par l'ensemble des éléments et des relations de dépendance qui la composent. Le modèle décrit ainsi les relations logico-sémantiques que peuvent entretenir les unités lexicales référant à une attitude avec d'autres éléments, lexicaux et grammaticaux, ainsi que l'impact que ces relations auront sur la polarité finale accordée à l'expression d'attitude. Le modèle comprend deux parties : la première décrit les règles sémantiques fondamentales, la seconde s'intéresse à la manière dont celles-ci sont appliquées dans la structure argumentale d'un certain nombre de verbes.

Groupe 1A	Appréciations & Jugements	
	NP-LinkVB-AdjP-PP	“The pitch is perfect for cricket” “He is modest about the extent to which it has brought him fame”
	NP-LinkVB-AdjP-InfClause	“Horses are pretty to look at” “You’re right to say that”
	Affects	
	NP-LinkVB-AdjP-toInfClause	“Benjamin had been rather overawed to meet one of the Billington family”
Groupe 1B	Affects, Appréciations & Jugements	
	It-LinkVB-AdjP-finite   nonFiniteClause	“it was reasonable to incur that expense” “It was wonderful talking to you the other day” “It was distressing to hear her talking like that”
	It-LinkVB-AdjP-SP + nonFiniteClause	“It was unnecessary of him to say it” “It was silly of/for them to do that” “It was touching of Pugh to place a sticker in the back of the hearse”
Groupe 2	Affects, Appréciations & Jugements	
	NP-OpinionVB-it-AdjP-thatClause	“i find it dreadful that a man in his position could do something”
	NP-OpinionVB-it-as-adjP	“They see it as beautiful”

TABLEAU 3.1 – Synthèse des patrons syntaxiques présentés dans (Bednarek, 2009)

Quatre règles fondamentales sont proposées (tableau 3.2) pour décrire les différents phénomènes linguistiques jouant sur la polarité. Ces règles opèrent à plusieurs niveaux : au niveau du syntagme lui-même (relation entre adjectif et adverbe, entre adjectif et nom), mais aussi à celui des relations de dépendance entre syntagmes (relation sujet-verbe, verbe-objet, etc).

À ces règles s’ajoute une analyse des structures verbales et des règles sémantiques qui s’appliquent lorsque le syntagme verbal et/ou l’un de ses arguments sont porteurs d’une attitude. Vingt-deux classes sont définies. Chacune d’elles se différencie par le rôle que ses membres jouent dans l’élaboration du sens d’une attitude. Parmi les classes définies, cinq impliquent des verbes référant à une attitude. Pour celles-ci, le modèle établit quelle règle sémantique s’applique en cas de conflit de polarité avec l’un des arguments. Pour les autres, le modèle définit l’impact que le sens du verbe aura sur l’attitude exprimée par l’un de ses arguments lorsque cela est le cas.

La question de la polarité est intrinsèquement liée à celle de la classification en sous-classes d’attitude. Deux interprétations sont proposées pour les expressions au sein desquelles le verbe

Inversion de polarité	<p>Un terme de négation en relation avec un segment référant à une attitude : <i>never</i> POS(<i>succeed</i>) = NEG(<i>never succeed</i>) ;</p> <p>Un adverbe de doute en relation avec un segment référant à une attitude : <i>scarcely</i> POS(<i>relax</i>) = NEG(<i>scarcely relax</i>) ;</p> <p>Mot référant à la diminution, au manque en relation avec un segment renvoyant à une attitude : adjective <i>reduced</i> POS(<i>enthusiasm</i>) = NEG(<i>reduced enthusiasm</i>)</p>
Agrégation	Dans un même syntagme, lorsque deux mots, référant à une attitude, sont en relation et que leurs polarités diffèrent, la polarité du mot ayant un statut de modifieur est assignée au groupe : POS( <i>beautiful</i> ) NEG( <i>fight</i> ) = POS( <i>beautiful fight</i> ) ;
Dominance	Lorsque la règle de dominance est appliquée, la polarité du verbe prévaut sur celle de l'objet : NEG( <i>to deceive</i> ) POS( <i>hopes</i> ) = NEG( <i>to deceive hopes</i> )
Propagation	La polarité du verbe se propage vers son objet : POS( <i>to admire</i> ) <i>his behaviour</i> = POS( <i>his behaviour</i> )
Neutralisation	La règle de neutralisation s'applique lorsqu'une préposition, exprimant le manque, l'absence, ou une conjonction de condition sont associées à un mot exprimant une attitude : <i>despite</i> NEG( <i>worries</i> ) = NEUT( <i>despite worries</i> )
Intensification	La règle d'intensification s'applique lorsqu'un modifieur accentue la polarité du mot véhiculant une attitude : <i>very</i> NEG( <i>sad</i> ) = NEG++( <i>very sad</i> )

TABLEAU 3.2 – Règles sémantiques de calcul de polarité (Neviarouskaya et al., 2010b)

réfère à une attitude : *interior perspective* et *exterior perspective*. La première est une interprétation littérale, qui se place du point de vue du sujet du verbe, expérienceur ou source de l'attitude exprimée, tandis que la seconde se place du côté du locuteur jouant ici le rôle d'évaluateur extérieur. Ainsi, selon l'interprétation choisie, une même phrase pourra être interprétée soit comme un affect soit comme un jugement, et la polarité pourra être inversée. Par exemple, pour la phrase (iii) ci-dessous, la première interprétation (*interior perspective*) permettra d'y lire l'expression d'un affect positif, dont l'expérienceur est le sujet du verbe (*He*) et la cible, l'objet (*mafia leader*). La seconde interprétation (*exterior perspective*) considérera la phrase comme l'expression du jugement négatif dont la source est le locuteur, et la cible, le sujet du procès *to admire mafia leader*. La première interprétation met en relief uniquement le système de valeur de l'expérienceur, tandis que la seconde ne le prend pas en compte et réfère au système de valeur en vigueur dans le contexte énonciatif.

(iii) “He admires mafia leader”

## 3.2 Formes syntaxiques, rôles sémantiques et calcul de la polarité

Les modèles présentés ci-dessus apportent des éléments essentiels pour comprendre la manière dont les attitudes sont exprimées sur le plan linguistique. Les patrons syntaxiques proposés par Martin and White (2005) et Bednarek (2009) sont pensés essentiellement pour résoudre des problèmes de classification : leur objectif est essentiellement de définir quel type d'énoncé réfère à quel type d'attitude. Les places occupées par la source et la cible ne sont pas analysées. Dans le travail de Neviarouskaya et al. (2010b), les règles sémantiques se concentrent sur le calcul de la polarité et sur celui de la catégorisation et ne considèrent que peu la forme syntaxique des énoncés. La question de la source et de la cible en est également absente. Cette section présente ainsi une extension et une fusion de ces deux modèles afin de constituer la base théorique nécessaire au développement de nos grammaires pour l'analyse phrastique des expressions d'attitude (détaillées en annexe F). Dans un premier temps, nous détaillerons les patrons syntaxiques que nous avons retenus ou ajoutés et questionnons la question de l'identification de la source et de la cible. Dans un second temps, nous présentons pour chaque niveau syntaxique la manière dont nous appliquons les règles de polarité.

### 3.2.1 Structures syntaxiques et identification de la source et de la cible

Pour décrire les patrons syntaxiques des expressions d'attitude, nous choisissons à la manière de Martin et White de partir du lexique. Pour les expressions analysées ici, nous considérons comme élément initial une unité lexicale référant à un processus axiologique ou affectif. Cet élément minimal de sens, associé aux relations logico-sémantiques de la phrase dans laquelle il s'intègre, formera le sens spécifique de l'attitude exprimée par l'énoncé. Selon Martin and White (2005) et Bednarek (2009), ces éléments minimaux de sens, utiles à l'expression d'attitudes, appartiennent à trois parties de discours différentes : noms, adjectifs et verbes. Néanmoins, dans la majorité des patrons syntaxiques qu'ils proposent, seuls les adjectifs sont considérés. Enfin de clarifier l'analyse des structures phrastiques, considérons, dans un premier temps, les syntagmes au sein desquels verbes, adjectifs, noms et adverbes référant à une attitude s'inscrivent (voir figure 3.1) :

- *Adj(Att)* : les adjectifs exprimant une attitude peuvent être employés dans des syntagmes adjectivaux (*SAdj*) ou des syntagmes nominaux (*SN*). Dans les premiers, ils constituent la tête du syntagme, dans les seconds, ils ont un statut de modifieur.
- Les *Nom(Att)* et *Vb(Att)* prennent place respectivement dans des syntagmes nominaux, *SN(Att)*, et verbaux, *SV(Att)*. Dans le syntagme nominal, la question de la cible se pose déjà : la relation de qualification que l'adjectif entretient avec le nom permet de considérer l'entité ou le concept dénoté par le nom comme la cible de l'attitude dénotée par l'adjectif. Néanmoins, cette pré-désignation de la cible est fréquemment relayée par une désignation plus forte définie au niveau de la phrase.

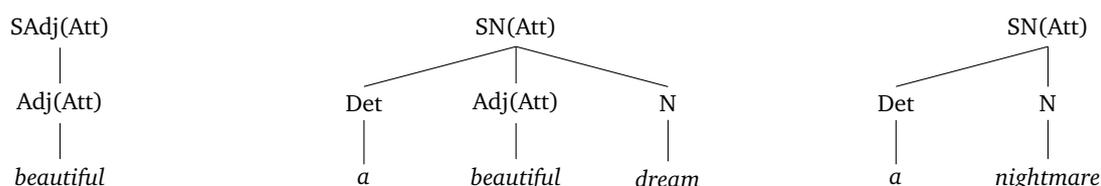


FIGURE 3.1 – Syntagmes adjectivaux et nominaux référant à une attitude

En plus des structures attributives décrites par Martin and White (2005) et Bednarek (2009), des structures prédicatives peuvent également accueillir les syntagmes décrits ci-dessus. À la dif-

férence des structures attributives, les structures prédictives expriment – grâce à leur prédicat – un procès ou un état dénoté soit par un nom soit par un verbe.

### 3.2.1.1 Structures attributives

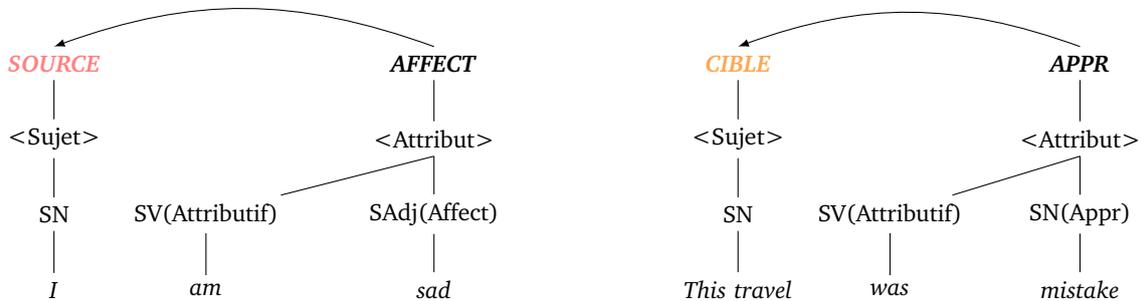


FIGURE 3.2 – Syntagme référant à une attitude en position attribut

Parmi les syntagmes référant à une attitude, seuls les syntagmes nominaux et adjectivaux (*SN(Att)* et *SAdj(Att)*) sont concernés par les structures attributives. Même si cela n’apparaît pas dans les patrons de Bednarek (2009) et Martin and White (2005), le syntagme nominal *SN(Att)* peut occuper soit la place d’<Attribut> soit celle de <Sujet>. L’identification de la cible et de la source dépendra alors de la fonction occupée par le syntagme porteur de l’attitude (*Synt(Att)*) et du type d’attitude à laquelle il réfère.

- **Attribut et affect** : lorsqu’il fait fonction d’attribut, s’il exprime un *affect*, le <Sujet> réfère alors à l’*Expérienceur* (ou *Source* dans le modèle de Martin and White (2005) (voir le premier schéma Figure 3.2).
- **Attribut et appréciation/jugement** : en revanche, s’il réfère à une appréciation ou jugement *SN(Att, Eval)*, le <Sujet> désigne alors la *Cible*. La source, dans ce type de structure, n’est pas verbalement exprimée, elle est donc associée au locuteur de l’énoncé (voir le second schéma figure 3.2).
- **Sujet** : lorsque *SN(Att)* occupe la fonction sujet et que l’attribut ne réfère à aucune attitude, la cible est alors à chercher *a minima* à l’intérieur même du syntagme *SN(Att)*. Dans la phrase, *This beautiful painting is a cubist*, la cible peut être assimilée à *painting* (voir Figure 3.3).

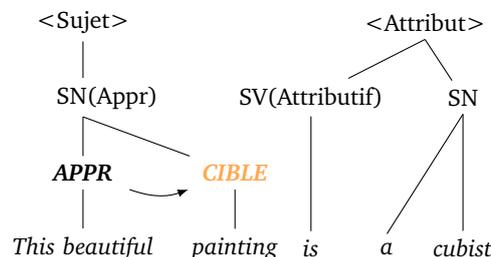


FIGURE 3.3 – Syntagme référant à une attitude en position sujet

**Ajout d’un syntagme prépositionnel ou d’une proposition** La structure de base <Sujet> <SV(Attributif)> <Attribut> peut être complétée par un syntagme prépositionnel (*SP*) et/ou une proposition infinitive (*INF*) ou participiale (*PART*). C’est ce type de structures que l’on retrouve dans la modélisation syntaxique proposée par Martin and White (2005) et Bednarek

(2009) (voir groupe 1A et 1B dans le tableau 3.1). L'ajout de tels compléments à la structure originelle peut avoir une incidence sur l'identification des rôles sémantiques propres à l'attitude exprimée :

1. La place de sujet n'est pas occupée par un pronom indéfini ("It") :
  - (i) si le *SAdj(Att)* ou le *SN(Att)* exprime un affect, le groupe rattaché à l'attribut (*SP*, *INF* ou *PART*) réfère alors au déclencheur de l'affect et le sujet, à son expérienceur (par exemple, "Benjamin had been rather overawed to meet one of the Billington family");
  - (ii) si le *SAdj(Att)* ou le *SN(Att)* exprime une appréciation ou un jugement, le groupe rattaché à l'attribut (*SP*, *INF* ou *PART*) ne joue ni le rôle de *cible* ni celui de *source* mais permet de contextualiser de l'évaluation et de resserrer le champs d'application de la propriété dénotée par l'attribut. Par exemple, dans la phrase "The pitch is perfect for cricket", l'attribution de la propriété *perfect* à *pitch* est présentée comme valable dans le contexte spécifique de *cricket*.
2. La place de sujet est occupée par un pronom indéfini : dans ce cas, le groupe rattaché à l'attribut (*SP*, *INF* ou *PART*) est en fait le référent du pronom placé en position sujet. Ainsi, dans la phrase "It was wonderful talking to you the other day", le sujet de référence de l'attribut est la proposition *talking to you the other day*. Ce type de structure nous permet d'évoquer la question de la coréférence. Ici, le groupe syntaxique de référence est rattaché à la structure globale de la phrase, sous forme de syntagme prépositionnel ou d'une proposition infinitive ou participiale. Néanmoins, dans certains cas, le référent peut apparaître dans les énoncés ou même les tours de parole qui précèdent.

**Structures à verbes d'opinion** Autre variante de la structure attributive de base, les structures intégrant des verbes d'opinion. Deux patrons parmi ceux de Martin and White (2005) et Bednarek (2009) intègrent à un verbe d'opinion (voir groupe B dans le tableau 3.1 et schémas 3.4 et 3.5). Bien que le verbe attributif soit absent, ces structures fonctionnent de manière identique à celles que nous venons d'aborder. La relation d'attribution de propriétés n'implique plus l'attribut et le sujet mais l'attribut et l'objet. Ces derniers ne constituent pas la prédication principale de l'énoncé mais la prédication seconde (Melis, 1988), la prédication principale étant occupée par le verbe de pensée ou d'opinion. Dans les deux cas, la présence d'un verbe d'opinion permet d'inscrire la source dans la structure syntaxique, celle-ci est alors exprimée par le sujet du verbe d'opinion. En revanche, pour la cible, deux cas de figure sont à envisager :

1. si l'attribut est complété par une proposition subordonnée ("I find it dreadful that a man in his position could do something like that") et que l'objet est un pronom indéfini, alors la cible est exprimée par la proposition subordonnée ;
2. si l'attribut n'est pas complété par une proposition subordonnée, la cible est exprimée par l'objet.

A ces deux patrons, peut également s'ajouter celui présenté Figure 3.12. Le fonctionnement est assez similaire, à la seule exception que le verbe attributif y est présent. Concernant l'attribution des rôles sémantiques, la source est exprimée par le sujet du verbe d'opinion et la cible par le sujet de l'attribut.

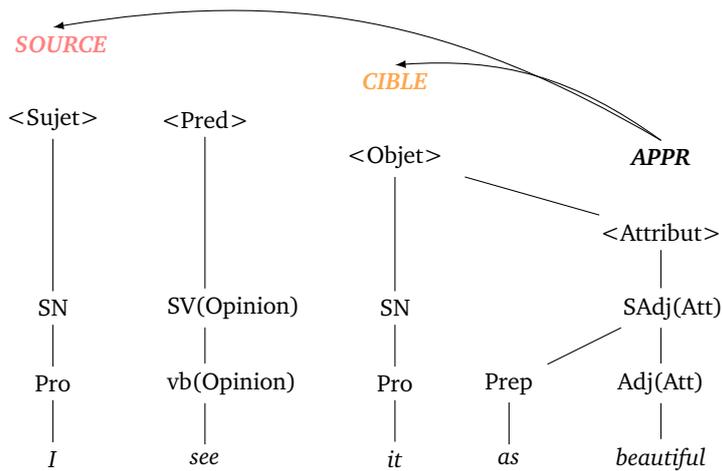


FIGURE 3.4 – Structure attributive à verbe d’opinion avec ellipse du verbe attributif

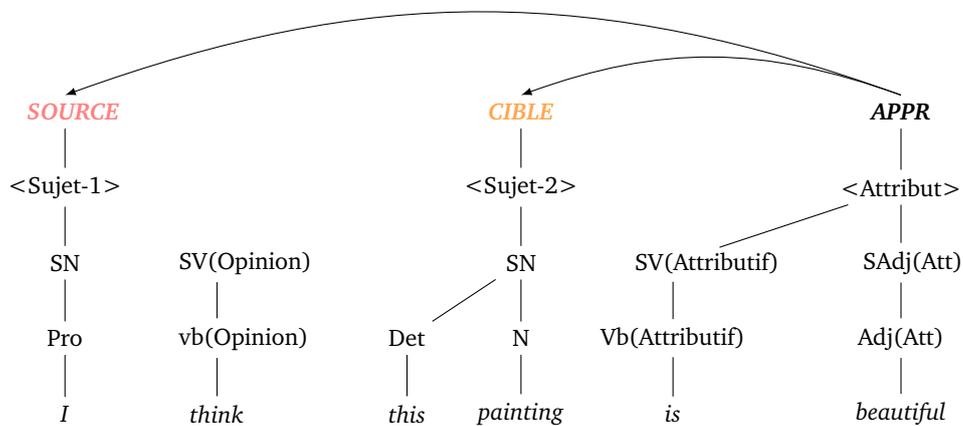


FIGURE 3.5 – Structure attributive à verbe d’opinion sans ellipse du verbe attributif

### Attribut de l’objet avec le verbe *to make*

NP-Make-NP-(Feel)\*-AdjP

“This painting makes me feel happy”

Ce patron, défini par Bednarek (2009), est également une structure attributive au sein de laquelle la propriété dénotée par l’attribut est attribuée à l’objet (*me* dans la phrase ci-dessus). Au sein de ce type d’énoncé, c’est l’objet du verbe *to make* qui réfère à la source, tandis que la cible est exprimée par le sujet du même verbe (*This painting*, dans la phrase ci-dessus). Une infinitive peut se rattacher à l’attribut de l’objet : dans ce type de configuration, si la place de sujet est occupée par un pronom, l’infinitive lui sert alors de référent et permet d’exprimer la cible. Par exemple, dans le “It made me feel happy to see you yesterday”, *me* est la source, tandis que la cible est la proposition infinitive *to see you yesterday*, référent du pronom *It*.

### 3.2.1.2 Structures prédictives

Les structures prédictives n'apparaissent pas dans le modèle de Martin and White (2005). À la différence des structures attributives, elles permettent de dénoter un procès impliquant un ou plusieurs actants. Le prédicat dénotant le procès peut être de forme verbale ou de forme nominale. Dans le cas d'une prédication nominale, un *SN(Att)* peut être employé pour en position sujet ou comme prédicat. Les *SN(Att)* employés comme prédicats réfèrent généralement à des affects ou des états émotionnels affectant le sujet. Dans ce type d'énoncé, la source est donc exprimée par le sujet du verbe support (Figure 3.6).

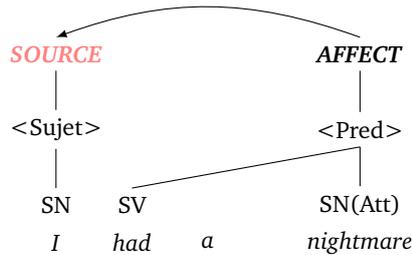


FIGURE 3.6 – Structure à prédicat nominal

Dans le cadre des structures prédictives dans lesquelles le prédicat est exprimé par un verbe, différentes configurations sont à envisager (voir figure 3.7). Tout d'abord, le *synt(Att)* peut être le syntagme verbale jouant le rôle de prédicat, comme par exemple dans la phrase "I like outdoors activities". Ensuite, il peut également jouer le rôle de l'un des arguments, comme dans la phrase "I saw a beautiful movie last night", où le *SN(Att)* joue le rôle d'objet du prédicat. Nous ne nous étendrons pas sur ce dernier cas dans ce chapitre. Nous considérons en effet que parmi les structures prédictives et verbales, ce sont celles où le verbe réfère à une attitude qui sont les plus à même de se manifester. Parmi celles-ci, nous distinguons celles dans lesquelles le sujet est la cause, le déclencheur du procès exprimé par le verbe et l'objet l'expérimenteur ("This movie annoys me"), de celles dans lesquelles, à l'inverse, le sujet réfère à l'individu expérimentant le procès et l'objet celui qui le provoque ("I like outdoors activities").

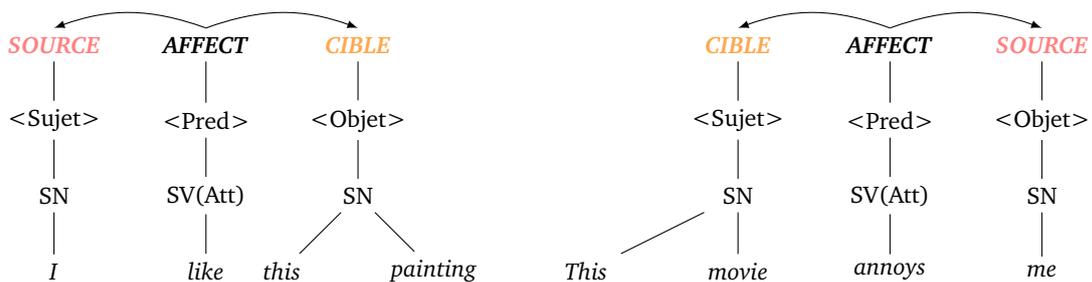


FIGURE 3.7 – Structures à prédicat verbal

### 3.2.2 Structures syntaxiques et règles de polarité

Concernant le calcul de la polarité, nous retenons les règles définies par Neviarouskaya et al. (2010b) et présentées dans la section 3.1.2. Ci-dessous nous décrivons la manière dont elles s'appliquent dans les structures présentées dans la section précédente.

### Règles de polarité dans le cadre syntagmatique

Dans le syntagme adjectival (Figure 3.8), deux règles peuvent s'appliquer.

- La règle d'intensification : elle s'applique lorsqu'un adverbe d'intensité porte sur l'adjectif.
- La règles d'inversion : elle s'applique lorsqu'un adverbe de doute porte sur l'adjectif.



FIGURE 3.8 – Règles de polarité dans le syntagme adjectival

Dans le cadre du syntagme nominal (voir figure 3.9), trois règles de calcul de polarité peuvent s'appliquer :

- la règle d'intensification s'applique en présence d'un intensifieur (déterminant, *such*, ou adjectif, *rising*) portant sur le nom ou sur l'adjectif qui le qualifie ;
- la règle d'agrégation s'applique lorsque l'adjectif portant sur le nom à une polarité différente de celle du nom. Dans ce cas, c'est la polarité de l'adjectif qui l'emporte ;
- la règle de neutralisation s'applique lorsqu'un déterminant marquant l'absence ou le manque (*no*, *any*) porte sur le nom ou l'adjectif qui le qualifie.

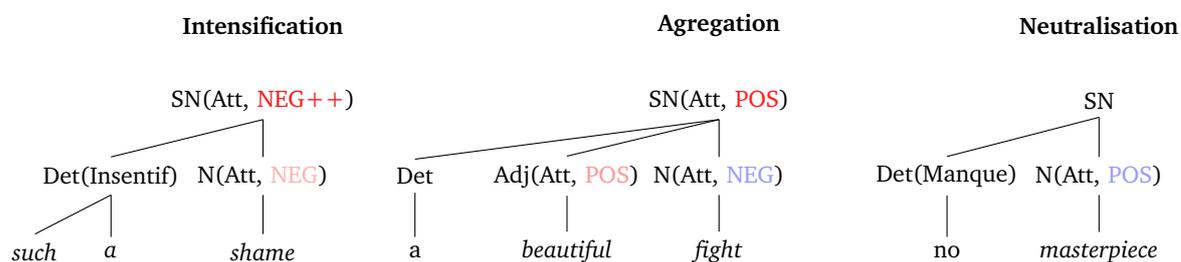


FIGURE 3.9 – Calcul de la polarité dans le syntagme nominal

Dans le cadre du syntagme verbal, la règle d'inversion s'applique lorsqu'une marque de négation (*not*, *never*, etc) porte sur le verbe.

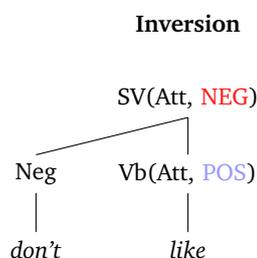


FIGURE 3.10 – Inversion de polarité dans le syntagme verbal

### Règles de polarité dans le cadre phrastique

Dans le cadre phrastique de type attributif, la polarité est déterminée par la polarité véhiculée par l'attribut du sujet ou de l'objet et par la nature affirmative ou négative du syntagme verbal attributif. Comme l'indique la figure ci-dessous, la négation du syntagme verbal attributif provoque l'inversion de la polarité véhiculée par l'attribut.

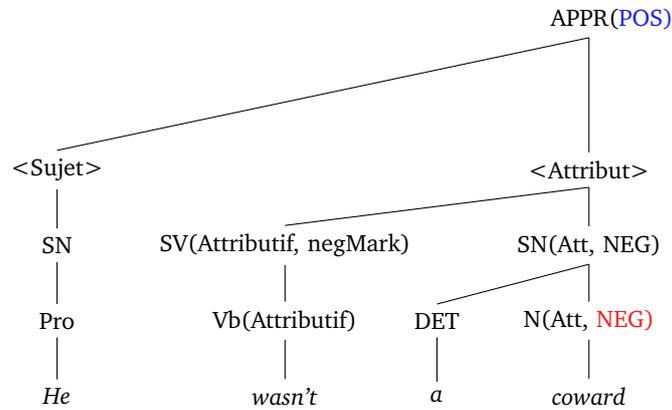


FIGURE 3.11 – Inversion de polarité dans le cadre d'une structure attributive

Précisons que l'effet de la négation peut osciller selon les locuteurs et les contextes. La négation peut impliquer la référence à une valeur soit contradictoire, soit contraire à celle exprimée par l'attribut nié. Ainsi la phrase, "It isn't a great book" peut avoir une polarité négative (contraire à la polarité positive de *great*) ou une polarité indéfinie se plaçant entre une polarité négative totale et une polarité neutre. Pour le français par exemple, cette question a pu être traitée en profondeur par la thèse de Zhang (2012b). Pour simplifier le modèle, nous avons choisi de considérer la négation d'un verbe attributif comme impliquant une inversion de la polarité véhiculée par l'attribut. Nous avons néanmoins considéré qu'une négation d'un verbe d'opinion avait un effet différent. Dans la phrase présentée dans la figure ci-dessous, la négation du verbe d'opinion semble davantage provoquer la neutralisation que l'inversion de la polarité.

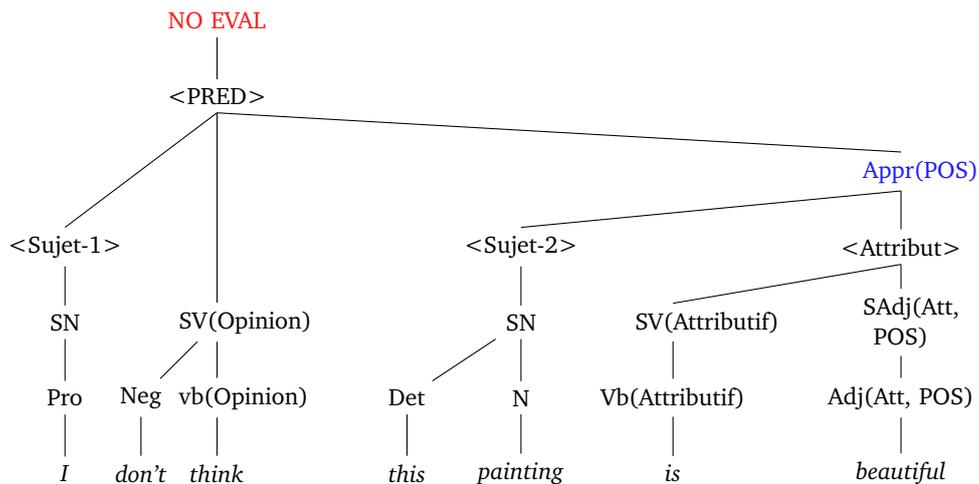


FIGURE 3.12 – Application de la règle de neutralisation dans le cadre d'une structure à verbe d'opinion

La règle de neutralisation s'applique également dans les structures prédictives nominales. Ainsi, la phrase "I didn't have nightmare last night" apparaît comme référent davantage à une absence de cauchemars (polarité neutre) qu'à la présence de beaux rêves (polarité positive).

### 3.3 Spécificités liées au contexte conversationnel

**Attitudes et parole orale** L'ensemble des structures détaillées dans ce chapitre a été identifié dans nos corpus d'étude. La nature orale de ces données ne semble pas altérer la structure profonde des énoncés : les utilisateurs exploitent des structures attributives ou prédicatives et les actants des attitudes exprimées occupent les mêmes positions logico-sémantiques. C'est davantage la structure de surface des énoncés qui semble subir des altérations. Nous avons ainsi rencontré des répétitions, des ellipses ou des perturbations dans l'organisation des éléments sur l'axe syntaxique. Par exemple, dans la phrase ci-dessous, on observe une structure attributive permettant de caractériser positivement le sujet *painting*. La structure profonde correspond à ce que nous avons pu analyser dans les sections précédentes. En revanche, la structure de surface présente des disfluences, que sont les répétitions et les hésitations.

(iv) hum, I think huum I think the painting is beautiful

Dans le cadre de notre thèse, nous avons choisi de ne pas créer des patrons dédiés à la gestion de ces disfluences. Nous considérons en effet que celles doivent faire l'objet d'une modélisation spécifique tel que cela a pu être proposé par C. Dutrey dans sa thèse de doctorat (Dutrey, 2014). En revanche, en l'absence d'un système de détection des disfluences facilement intégrable à notre chaîne de traitement, nous avons choisi d'ajouter une certaine flexibilité à l'intérieur de nos grammaires (décrites chapitre 6) afin que celles-ci puissent gérer *a minima* certaines répétitions et certaines ellipses.

**Small-Talk vs négociation** L'observation du corpus de négociation et du corpus Semaine n'a pas montré de différence significative ni dans les formes syntaxiques utilisées ni dans les mécanismes sémantiques. Les structures et règles détaillées dans les sections précédentes sont présentes dans chacun des deux corpus. Dans la partie suivante (chapitre 6), d'autres types d'énoncés seront pris en compte par le système de détection : des énoncés volitifs permettant à l'utilisateur d'exprimer souhait et volonté dans le cadre notamment de la négociation. Ces énoncés ne sont pas à proprement parler des attitudes. Ils sont pertinents dans le cadre des *likes* et des *dislikes* de l'utilisateur telle que l'abordera le chapitre 6, mais sont particulièrement liés au contexte de la négociation. C'est pour cette raison qu'ils n'apparaissent pas dans le présent chapitre. Dans les figures 3.13 et 4.3, nous présentons des extraits du corpus Semaine et du corpus de négociation. Dans chacun d'eux, les expressions d'attitude correspondant aux structures présentées dans ce chapitre apparaissent en rouge.

Dans la session du corpus Semaine (figure 3.13), trois structures prédicatives intégrant le verbe *like* peuvent être identifiées (2, 3 et 8). Elles ont pour source l'utilisateur. La cible varie pour chacune d'elles : *to go away for a holiday in easter*, pour l'expression 2, *to go to Santorini*, pour la 3 et *this weather* pour la 8. En ce qui concerne les règles de polarité, on note au sein de l'expression 7 une négation portant sur le verbe *to like* et déclenchant une inversion de polarité. Toujours en ce qui concerne les structures prédicatives, une expression avec un prédicat nominal exprimant un affect (expression 1). La source est bien l'utilisateur (pronom personnel de première personne placé en position sujet du verbe support) et le déclencheur de cet affect négatif n'est autre que la semaine vécue par l'utilisateur. Enfin, quatre expressions ont une structure attributive (4, 5, 6 et 7). Deux d'entre elles sont du type <SUJET> <VbAttributif> <ATTRIBUT> (4 et 6). Dans la 4, à la place du sujet de l'attribut, se trouve un pronom référant à *Santorini* dans le tour de parole précédent. Et dans la 7, la place de sujet est également occupée par un pronom mais dont le référent se trouve dans le même tour de parole, *islands*. Enfin, les expressions 5 et 7 montrent une structure intégrant le verbe *to make*. La cible occupe la place de sujet du verbe : *islands*, pour la 5 et, *a bit of sun* pour la 7.

- 1 O : Hi my name is Poppy  
 U : Hello  
 O : I'm the eternal optimist  
 U : Well you'd need to be because **i've had an awful week**  
 O : An awful week?  
 U : Terrible  
 O : That's not good, that's not good to hear em well things will normally get better, they normally eh look up can you think of something that might make you happy
- 2 U : No, at the minute it's pretty hard but i could always think about holidays **I'd like to go away for a holiday in easter** but I'm only in the month of january so no point thinking about that  
 O : Aha. Where do you think you'll go can you tell me about your holiday?  
 U : **I'd like to go to San Torrini**  
 O : Aha. Can you tell me about that?  
 U : Ah **it's a lovely island in a lovely place** but I suspect my husband's arranged to do some work and I'll not be able to get there so that doesn't really help me
- 3 O : Aha but if you think about it, it makes you happy?  
 U : Yes but I might just end up going to donegal instead or galway or somewhere  
 O : That's good and does do they make you happy?  
 U : I haven't been there for a long time but **islands generally make me feel happy they kinda like nice peaceful restful places. But a bit of sun would also make me feel happy**
- 4 O : That's good  
 U : **I don't like this weather**  
 O : Ok and eh so the sun you so you find that uplifting  
 U : Oh yes, very
- 5  
 6  
 7  
 8

FIGURE 3.13 – Attitudes dans le tour de parole de l'utilisateur. Extrait du corpus Semaine

Dans la session de négociation (Figure 4.3), deux expressions d'attitude peuvent être identifiées. Il s'agit de structures prédicatives dont la cible, pour l'expression 9, est *all these items*, et pour la 10, *the picture*.

- 9 A : So, what are you most interested in?  
 U : uh **I'm pretty much equally interested in all these items**  
 A : Oh really?  
 U : They're they're all of uh equal value to me  
 A : huh I think you must have a preference for one thing more than the others  
 U : **I probably prefer the picture**
- 10 A : Ok. Do you like the painting?  
 U : Yes  
 A : Did you want any of the records?  
 U : yeah  
 A : Alright well what do you propose?  
 U : um I take one more of these and you can have both lamps  
 A : Well, that's an interesting idea. let's move the items to show what we're talking about  
 U : there you go

FIGURE 3.14 – Attitudes dans le tour de parole de l'utilisateur. Extrait du corpus de négociation

### 3.4 Conclusion du chapitre 3

Dans ce chapitre, nous avons fourni une description des formes syntaxiques et des mécanismes sémantiques des expressions d'attitude dans le tour de parole de l'utilisateur. Pour cela, nous nous sommes appuyée sur les patrons syntaxiques et les règles sémantiques existant dans la littérature. Afin de faciliter le développement des grammaires intégrées au système (voir chapitre 6), nous avons cherché à compléter ces ressources avec des informations supplémentaires et à les fusionner. Concernant les patrons syntaxiques, nous avons analysé leur structure profonde afin de mettre au jour les mécanismes d'attribution des rôles sémantiques de *source* et de *cible* dans les différentes expressions d'attitude. Partant des syntagmes pour aller jusqu'aux structures phrastiques, nous avons ainsi identifié parmi les différents groupes ou unités lexicales composant une expression d'attitude, ceux permettant de référer à la cible et ceux référant à la source. Cette analyse nous a également permis de faire émerger certaines structures absentes du modèle de Martin and White (2005) et de Bednarek (2009). Concernant les règles sémantiques, nous avons repris une partie des règles proposées par Neviarouskaya et al. (2010b) et analysé la manière dont elles s'appliquaient dans les structures syntagmatiques et phrastiques précédemment étudiées. Nous avons ainsi observé les effets de la compositionnalité sémantique à l'œuvre dans la phrase pour le calcul de la polarité. Enfin, nous avons conclu ce chapitre par une exemplification sur le corpus Semaine et sur le corpus de négociation montrant la similarité des formes employées dans les deux contextes.



## Chapitre 4

# Les attitudes à l'échelle de la paire adjacente

### Sommaire

---

4.1	Les attitudes dans le cadre collaboratif du dialogue . . . . .	47
4.1.1	Dialogue et contributions . . . . .	47
4.1.2	Les Attitudes, des Actes Participatifs . . . . .	48
4.1.3	Attitudes et Contributions dans les Conversations Humain-Agent . . . .	49
4.2	Expressions d'attitude strictement collaboratives . . . . .	49
4.2.1	Réponses aux questions fermées . . . . .	50
4.2.2	Réponses aux questions ouvertes . . . . .	51
4.2.3	Assertion d'une attitude . . . . .	53
4.3	Formalisation des attitudes strictement collaboratives . . . . .	54
4.4	<i>Small-Talk</i> vs négociation . . . . .	54
4.5	Conclusion du Chapitre 4 . . . . .	56

---

DANS le chapitre précédent, nous avons analysé les attitudes telles qu'elles s'expriment à l'échelle du tour de parole. Afin d'améliorer la compréhension du phénomène en contexte conversationnel, il est nécessaire de dépasser le cadre du tour de parole et d'observer la distribution et le fonctionnement des tours de parole entre eux. Pour cette raison, le cadre d'étude du présent chapitre est élargi à celui de la paire adjacente Sacks et al. (1974). Notre objectif est d'analyser le fonctionnement des expressions d'attitude à l'intérieur de ce cadre et de montrer que certaines d'entre elles ne peuvent être analysées sans une prise en compte conjointe de l'apport sémantique des deux énoncés qui composent la paire adjacente. Ces expressions, que nous appelons *expressions d'attitude strictement collaboratives* (EASC), se distinguent de celles étudiées dans le précédent chapitre, que nous nommons *expressions simples d'attitude*, en ce que l'attribution des valeurs aux attributs *polarité*, *cible* et *source* repose exclusivement sur la collaboration dialogique des locuteurs (dans notre cas, l'agent et l'utilisateur). Dans un premier temps, nous interrogeons la notion de contribution dialogique telle que décrite par Clark and Schaefer (1989), qui permet d'envisager la conversation non comme une simple suite d'énoncés mais comme un travail collaboratif, dans lequel chaque énoncé implique la participation des deux interlocuteurs, du côté de la production et de la réception. Pour les auteurs, l'unité de base de cette contribution dialogique est la paire adjacente. La première section de ce chapitre est pour nous l'occasion de questionner le caractère participatif et collaboratif des attitudes. Mettant en valeur un sous-phénomène particulier à la nature contributive des attitudes en conversation, celui des EASC, nous nous efforçons, dans un deuxième temps, d'analyser le fonctionnement sémantique de ces expressions spécifiques et d'en proposer une formalisation. Dans un troisième temps, nous procédons à une exemplification des phénomènes, décrits dans le chapitre, sur deux sessions issues de nos corpus de travail.

Publications associées à ce chapitre :

- Langlet, C. and Clavel, C. (2014b). Modelling user's attitudinal reactions to the agent utterances: focus on the verbal content. In *5th International Workshop on Corpora for Research on Emotion, Sentiment & Social Signals (ES3 2014)*, Reykjavik, Iceland
- Langlet, C. and Clavel, C. (2014a). Modélisation des questions de l'agent pour l'analyse des affects, jugements et appréciations de l'utilisateur dans les interactions humain-agent. In *Proceedings of TALN 2014 (Volume 2: Short Papers)*, pages 479–484, Marseille, France. Association pour le Traitement Automatique des Langues

## 4.1 Les attitudes dans le cadre collaboratif du dialogue

En conversation, les énoncés ne sont pas indépendants les uns des autres, ils s'inscrivent dans un travail coopératif entre les locuteurs. S'interrogeant sur la manière dont les locuteurs participent au dialogue, les auteurs décrivent les contributions dialogiques comme faisant partie d'actes collectifs. Dans cette section, nous présentons une synthèse du modèle de Clark and Schaefer (1989) et interrogeons la nature participative et contributive des expressions d'attitude.

### 4.1.1 Dialogue et contributions

Dans leur article, Clark and Schaefer (1989) s'interrogent sur la manière dont les locuteurs contribuent au discours. Le simple fait de produire le bon énoncé au bon moment n'est pas, selon eux, un critère suffisant. Ainsi, plus que des séquences d'énoncés produits tour par tour, les conversations doivent être appréhendées comme le fruit d'un travail collaboratif entre les différents participants qui, lorsqu'ils prennent la parole, doivent prendre en compte l'autre et s'assurer que celui-ci reçoit et comprend le message. À chaque tour de parole, qu'il s'agisse de celui qui parle ou de celui qui écoute, chacun a son rôle à jouer dans la progression de la conversation : le locuteur doit s'efforcer de préciser le contenu de sa contribution et de rendre celle-ci intelligible ; l'interlocuteur, quant à lui, s'efforce d'en enregistrer le contenu. Beaucoup d'unités de conversation – mots, phrases, etc – sont produites par un locuteur agissant de manière autonome. Mais les contributions sont élaborées par un locuteur agissant de manière collective. Il n'y a contribution que si le contributeur et son partenaire coordonnent leurs actions dans le même sens. Chaque contribution doit ainsi être envisagée comme un acte individuel participant à un acte collectif. Lorsque le locuteur A produit son énoncé, trois actes sont alors impliqués : la production de l'énoncé de A, l'intégration du contenu de l'énoncé par l'interlocuteur B et enfin l'acte collectif de A et de B qui ajoutent le contenu de l'énoncé à leurs connaissances partagées. Ainsi, quand Ann dit à Bob qu'elle est occupée, elle effectue un acte individuel, une assertion, qui contribue au dialogue. Mais cet acte n'est produit que comme la part d'un acte collective dans lequel Bob a également son rôle à jouer. Ainsi collectivement Ann et Bob ajoutent le contenu *Ann est occupée* à leurs connaissances partagées et cet acte collectif repose sur l'acte individuel de Ann qui formule une assertion et de Bob qui enregistre l'énoncé de Ann comme une contribution au dialogue.

D'un point de vue conceptuel, une contribution s'effectue en deux phases : une *phase de présentation* et une *phase d'acceptation*. Dans la *phase de présentation*, le locuteur A présente au locuteur B un énoncé *e* qu'il doit prendre en compte. En l'absence d'indication contraire, le locuteur A suppose que le locuteur B a compris ce que A signifiait en produisant l'énoncé *e*. Dans la phase d'acceptation, B accepte l'énoncé *e* en donnant la preuve *p* qu'il pense avoir compris ce que signifiait l'énoncé *e*. Là encore, par cette action, il suppose que A lui donnera la preuve *e* qu'il croit que B l'a compris. Si ces deux phases sont effectuées correctement, alors A et B considéreront qu'ils sont parvenus à la connaissance partagée que B comprend ce que A signifiait en produisant son énoncé. Ces différentes phases sont celles qui sont nécessaires pour que A puisse contribuer au dialogue. Ajoutons que B peut utiliser différents types de preuve pour montrer à A qu'il a compris ce qu'il disait : soit en continuant à prêter attention à ce que dit A, soit en le remerciant par un signe de tête ou par une interjection de type "hum hum", soit en produisant un énoncé qui, à son tour, sera le point de départ d'une nouvelle contribution, soit encore en répétant tout ou une partie de ce que A a dit. Bien évidemment, le type et le nombre de preuves disponibles dépendent de la nature de l'interaction. Dans une conversation face à face, les locuteurs pourront avoir recours à des gestes, des regards ou des expressions faciales, tout autant d'indices qui ne fonctionneront pas dans le cadre d'une conversation téléphonique par exemple.

Ces contributions à la progression de la conversation coïncident le plus souvent au tour de parole et la plupart des tours de parole s'organisent en paires adjacentes. Une paire adjacente

se composent de deux tours de paroles. La forme et le contenu du second dépendra du type et du contenu du premier (“two ordered utterances, the first and second pair parts, produced by two different speakers. The two parts come in types that specify which is to come first and which second; the form and content of the second part”, (Clark and Schaefer, 1989)). Pour que la progression de la conversation suive son cours et que soit maintenue la cohérence discursive, la seconde partie doit donc remplir un critère de pertinence et d’acceptabilité.

Type of Adjacency Pair		Example	
First Part	Second Part	A's Utterance	B's Response
Question	Answer	Where is Connie?	At the store.
Request	Compliance/Refusal	Please pass the horseradish.	[B passes horseradish.]
Request	Acceptance/Rejection	Please pass the horseradish.	Okay.
Proposal	Acceptance/Rejection	Here is your change.	[B takes it.]
Offer	Acceptance/Rejection	Would you like some coffee?	Yes, thanks.
Invitation	Acceptance/Rejection	Come to dinner Sunday	Okay.
Apology	Acceptance/Rejection	Sorry.	Oh, that's all right.
Thanks	Acceptance/Rejection	Thank you.	You're welcome.
Assessment	Agreement/Disagreement	That film was terrible	Yes, it was.
Compliment	Agreement/Disagreement	Your new coat is beautiful	Yes, it's nice.
Summons	Answer	Hey, Ben	Yes?
Greetings	Greetings	Hi, Ben.	Hi, Ann.
Farewell	Farewell	Goodbye.	Goodbye.

FIGURE 4.1 – Typologie des paires adjacentes (Clark and Schaefer, 1989)

#### 4.1.2 Les Attitudes, des Actes Participatifs

En ce qui concerne les attitudes, selon la définition qu’en donnent Martin and White (2005), nous pouvons poser l’hypothèse qu’elles peuvent être soit de nature participative soit de nature autonome. Si certaines attitudes exprimées en conversation peuvent être considérées comme des actes individuels autonomes (le locuteur peut exprimer un jugement qui n’a pas pour vocation à contribuer au dialogue, mais qui se présente plus comme un commentaire), la plupart d’entre elles doivent être envisagées comme des actes participatifs émanant de la volonté du locuteur d’apporter sa contribution au discours et d’ajouter son avis ou ses goûts aux connaissances partagées. Deux configurations peuvent être envisagées : (i) soit l’attitude occupe la première partie de la paire adjacente ; (ii) soit l’attitude est exprimée dans la seconde partie. Puisque notre travail concerne la détection des attitudes de l’utilisateur et que celui-ci est moins en position d’ouvrir des paires adjacentes que l’agent, nous ne nous intéresserons qu’à la seconde configuration.

Dans l’exemple (i), B montre qu’il a compris ce que A signifie en exprimant un “yeah” et en exprimant à son tour une attitude. L’attitude exprimée au même titre que l’interjection d’acceptation *yeah* constitue donc la preuve de compréhension de l’énoncé de A. B respecte le critère de pertinence puisqu’il s’exprime sur ses goûts ainsi qu’il y est invité et que la cible concernée entretient un lien ontologique direct avec celle désignée dans l’énoncé de A. Agissant ainsi, il permet de valider l’énoncé de A comme contribution au dialogue, mais il fournit également le point de départ d’une prochaine contribution : il constitue à la fois la phase d’acceptation de l’énoncé A et la phase de présentation de la contribution B.

Exemple (i)

A : "I like outdoors activities"

B : "Yeah, I don't like being indoors"

### 4.1.3 Attitudes et Contributions dans les Conversations Humain-Agent

Du point de vue de la détection des attitudes dans les conversations humain-agent, plusieurs éléments sont à prendre en considération.

- Comme l'expliquent Clark and Schaefer (1989), le processus d'ajout de contenu aux connaissances communes construites par la conversation ainsi que la manière d'ajouter du contenu varient selon le type de conversation. Dans une conversation humain-agent, le contenu verbal de l'agent est scripté et repose sur un scénario d'interaction fixant les objectifs communicationnels de ce dernier. Les contributions de l'agent dépendent du bon déroulement du scénario et donc également de la capacité de l'utilisateur à participer activement à cette interaction. Si lorsque l'agent pose une question, l'utilisateur ne répond pas, il y a absence de contribution et la conversation ne peut avancer.
- Concernant les attitudes, l'agent doit pouvoir déterminer si elles s'inscrivent bien dans un travail collaboratif : l'utilisateur exprime-t-il une attitude parce que cela lui a été demandé ou suggéré ou, au contraire, est-ce l'expression autonome d'une attitude sans lien direct avec le contenu de la conversation ?
- Enfin, il est nécessaire de noter que certaines expressions d'attitude ne peuvent être interprétées que si cet aspect collaboratif est pris en compte. Ainsi dans l'exemple (ii) tiré du corpus Semaine, l'expression d'attitude repose sur les deux énoncés et ne se fait que par le jeu de question-réponse. À la différence de l'exemple (i) donné précédemment, le sens de l'attitude exprimée ne peut être connu que si une analyse de ce travail collaboratif est effectuée : tous les éléments de sens sont contenus dans la question et la réponse de l'utilisateur permet de les confirmer. En ce sens, ce type d'expression, que nous nommerons *expressions d'attitude strictement collaboratives* (EASC), propre au contexte dialogique et conversationnel, n'est intelligible que par la prise en compte des deux énoncés. Pour ce type d'attitude, le contexte de la paire adjacente et les mécanismes pragmatiques et sémantiques qu'elle impliquent se présentent comme un contexte d'analyse indispensable.

Exemple (ii)

Agent : "Do you like the sun?"

User : "Yes, I do"

## 4.2 Expressions d'attitude strictement collaboratives

Les expressions d'attitudes strictement collaboratives (EASC) reposent sur la combinaison des deux énoncés qui composent la paire adjacente. L'agent ouvre la paire adjacente en formulant un énoncé référant à une attitude, selon la réponse de l'utilisateur, une attitude dont il est la source sera actualisée. La typologie des paires adjacentes (figure 4.1) proposée par Clark and Schaefer (1989) offre un point de départ à l'analyse de ce type de structures. Parmi les énoncés proposés en première partie de la paire adjacente quatre d'entre eux sont pertinents pour notre propos : les questions, les requêtes, les assertions et les compliments. Dans un premier temps, notons que deux d'entre eux sont décrits comme référant à des attitudes. Les deux exemples ci-dessous (Tableau 4.1) présentent une expression d'appréciation. Ces énoncés sont donc le lieu pour le locuteur d'exprimer un avis ou une opinion. La différence réside dans la nature de la cible. Dans le second cas, la cible est l'interlocuteur tandis que dans le premier cas, la cible est extérieure à la conversation. Cette différence de nature des cibles n'a aucune conséquence quant à l'intérêt que ce type de phrase représente pour nous. Dans les deux cas, si l'agent formule ce

type d'énoncé, il y a possibilité que l'utilisateur exprime un accord ou un désaccord, impliquant l'expression d'une attitude dont il sera à son tour la source.

Assertion – Accord/Désaccord	Compliment – Accord/Désaccord :
A “That film was terrible” B “Yes, it was”	A “Your new coat is beautiful” B “Yes, it’s nice”

TABLEAU 4.1 – Exemple d'assertion et de compliment dans des paires adjacentes

À la différence des assertions et des compliments, les questions et les requêtes ne réfèrent pas nécessairement à des attitudes, mais cette possibilité reste néanmoins ouverte. Ainsi, une question peut être une simple demande d'information concernant l'état affectif, soit de l'utilisateur (“Are you happy today?”), soit d'une source quelconque (“Did he enjoy Christmas?”). Dans le premier cas, la simple expression d'accord ou de désaccord permet l'actualisation d'une attitude de l'utilisateur. Dans le second cas, en revanche, l'accord ou le désaccord à eux seuls ne déclenchent pas ce type d'expression, ils permettent seulement de vérifier l'existence d'une attitude dont la source est un tiers non impliqué dans la conversation.

L'objectif de l'analyse ci-dessous est de **comprendre la nature de l'implication de l'agent et de l'utilisateur dans l'attribution de valeurs aux attributs source, cible et polarité relatifs à l'attitude exprimée**. Les énoncés y sont décrits du point de vue de l'acte de dialogue qu'ils produisent et de celui de leur contenu propositionnel. Les attributs sémantiques constitutifs de l'attitude exprimée sont considérés comme composant du contenu propositionnel.

À ce point de l'analyse, il nous paraît indispensable de distinguer le rôle de locuteur de celui de source. Le locuteur est celui ou celle qui parle et la source, celui ou celle pour qui l'attitude est vraie qui en assume la teneur. Il est ainsi possible de référer à une attitude dont la source est une personne autre. De la même manière qu'un locuteur peut rapporter les propos d'autrui, il lui est possible d'en rapporter les attitudes. De fait, la source d'une attitude n'est pas nécessairement celui qui l'exprime verbalement. Pour cela, il est important de préciser que lorsque nous utilisons le terme *expression d'attitude*, nous désignons ainsi tout énoncé permettant de référer à une attitude dont la source n'est pas nécessairement le locuteur.

Enfin parce que l'analyse considère les énoncés du point de vue de l'acte de dialogue qu'ils produisent, elle porte des considérations pragmatiques. Ce que nous questionnons ici, ce sont les effets des énoncés. Néanmoins, nous considérons que ces énoncés s'incarnent dans des structures syntaxiques et sémantiques nécessaires à leur expression. Ainsi, si nous mettons l'accent ici sur des phénomènes pragmatiques, nous ne les considérons pas comme étant dissociés de problématiques syntaxiques.

#### 4.2.1 Réponses aux questions fermées

Les **questions fermées** invitent l'interlocuteur – ici l'utilisateur – à se positionner en validant ou en invalidant leur contenu propositionnel. Lorsque le contenu propositionnel concerne une attitude, la valeur de la source et celle de la cible sont prédéfinies et une orientation est donnée à la polarité. Ainsi, dans l'exemple (1) du tableau 4.2, le contenu propositionnel réfère à une évaluation potentielle non encore actualisée dont la source serait l'utilisateur et la cible “outdoors activities”. La question oriente cette attitude potentielle vers une attitude positive. L'orientation de la polarité est déterminée par des indices morpho-syntaxiques, ici le verbe dans une forme simple non négative. Dans la phrase (2) du même tableau 4.2, l'orientation négative est définie par l'usage d'une négation portant sur le verbe *love*. Le contenu propositionnel en question porte donc sur l'existence d'une appréciation négative de l'utilisateur (source) sur Picasso (cible).

Les caractéristiques sémantiques relatives à la cible, la source et la polarité, telles qu'elles sont présentées dans la question, sont essentielles, puisque c'est sur leur base que pourront être

Orientation positive de la polarité	Orientation négative de la polarité
(1) "Do you like outdoors activities?"	(2) "Really, you don't love" (3) "You really hate Picasso?"

TABLEAU 4.2 – Orientation de polarité dans les questions ouvertes

définies celles de l'attitude finale. C'est la réponse de l'utilisateur qui, ensuite, en validant ou en invalidant leurs valeurs, pourra actualiser l'attitude en question. L'effet de la réponse dépend en partie de la formulation de la question. Sans marque de négation, une réponse sous forme d'accord permet de conserver la valeur de la polarité telle que la prédéfinissait la question. Le tableau 4.3 présente des questions avec ou sans marque de négation et les réponses de type accord/désaccord qui leur sont associées. Pour la question (1), l'orientation positive de la polarité est conservée pour la réponse (1a). Une réponse de type *yes* équivaut alors à un accord. Il en va de même pour la question (3), dont la polarité négative est maintenue par la réponse (3a). En revanche quand, une négation porte sur le verbe (2), une réponse de type *yes* (2a) équivaut à un désaccord et actualise une attitude dont la polarité est l'inverse de celle prédéfinie par la question. Au contraire, une réponse de type *no* exprime un accord, la polarité est alors conservée. Cela permet de vérifier que syntaxe et pragmatique s'articulent et ne sont pas dissociées l'une de l'autre.

Question	Réponse Type <i>Yes</i>	Réponse Type <i>No</i>
(1) "Do you like Picasso?"	(1a) "Yes, I do" = Accord (User likes Picasso)	(1b) No, I don't" = Désaccord (User doesn't like Picasso)
(2) "You don't like Picasso?"	(2a) "Yes, I like him" = Désaccord (User like Picasso)	(2b) "No, I don't like his paintings" = Accord (User doesn't like Picasso)
(3) "Do you hate Picasso?"	(3a) "Yes, I do" = Accord (User hates Picasso)	(3b) "No, that isn't true" = Désaccord (User likes Picasso)

TABLEAU 4.3 – Accord et désaccord à une question fermée

Les questions du tableau 4.4 ont une structure de surface interrogative et leur structure sous-jacente est de type <source> <SV(Att)> <cible> présentée dans le chapitre précédent (chapitre 3). Cette structure peut varier et les différents types décrits dans le chapitre 3 peuvent être employés. De même, toutes les catégories d'attitudes sont concernées et les effets des différents types de réponses restent les mêmes.

#### 4.2.2 Réponses aux questions ouvertes

À la différence des questions fermées, les **questions ouvertes** n'invitent pas l'utilisateur à confirmer ou infirmer un contenu propositionnel, mais à le compléter. La réponse attendue n'aura donc pas la forme d'un accord ou d'un désaccord. Si les questions fermées prédéfinissent la valeur de la cible et de la source et orientent la valeur de la polarité, les questions ouvertes laissent l'un de ces attributs sans valeur. C'est sur cet attribut que portera la question et c'est sa nature qui décidera du type d'adverbe interrogatif utilisé. Les trois caractéristiques principales d'une attitude – polarité, cible, source – peuvent ainsi être questionnées. Ainsi, la question (1a) (Tableau 4.5) prédéfinit les valeurs de la source et de la polarité : l'agent s'adresse à l'utilisateur

Question	Réponse Type Yes	Réponse Type No
Did you have good trip ?	“Yes, I do Appreciation(pos, user, trip)	“No I don’t” Appreciation(neg, user, trip)
Did this movie annoy you ?	“Yes, it does Affect(neg, user, movie-Coref)	“No it doesn’t” Appr(neg, user, movie-Coref)
Do you think he is a brave man ?	“Yes I do” Judgment(pos, user, manCoref)	“No I don’t” Judgment(neg, user, manCoref)
Does it make you happy ?	“Yes, it does” Affect(pos, user, Coref)	“No it doesn’t” Affect(neg, user, Coref)

TABLEAU 4.4 – Questions fermées portant sur différents types d'attitude

et l'engage à exprimer une attitude positive, comme l'indique l'adjectif *favorite*, dont il est la source et dont la nature de la cible est à définir au sein d'une catégorie donnée “present”. En revanche, la question (2a) interroge la polarité : l'utilisateur est invité à définir un type d'affect dont la polarité peut varier du positif au négatif. Enfin, pour la question (3a), c'est l'identité de la source qui est en question : la réponse doit définir une source pour un affect positif *love* portant sur la cible *Picasso*.

Attribut Fondamental	Question	Réponse
Cible	(1a) “What was your favorite present?”	(1b) “A Playstation 4”
Polarité	(2a) “How do you feel today?”	(2b) “Good, I am good”
Source	(3a) “Who loves Picasso?”	(3b) “Dummies”

TABLEAU 4.5 – Questions ouvertes portant sur une caractéristique fondamentale

Les questions ouvertes peuvent concerner des caractéristiques circonstancielles et non fondamentales à la compréhension de l'attitude exprimée. Ces caractéristiques indiquent dans quelles *circonstances* l'attitude exprimée est valable. Il peut s'agir de circonstances de temps, de lieu, de causalité. L'élément circonstanciel, sur lequel porte la question, définit également le type d'adverbe interrogatif utilisé (*where*, pour le lieu, *when*, pour le temps, *why*, pour la cause, etc.). Ainsi, la question (1a) (tableau 4.6) interroge les circonstances de lieu de l'attitude exprimée par le syntagme nominal *best birthday*. La question (2a) interroge sur les circonstances temporelles de l'attitude *liking surrealism*. Et enfin, la question (3a) interroge les circonstances de causalité.

Attribut Circonstanciel	Question	Réponse
Localisation	(1a) “Where did you pass your best birthday?”	(1b) “ at San Diego”
Temporalité	(2a) “When did you start liking Surrealism? ”	(2b) “when I was in High School”
Causes	(3a) “Why do you love Picasso?”	(3b) “because his work is prodigious”

TABLEAU 4.6 – Questions ouvertes portant sur une caractéristique circonstancielle

### 4.2.3 Assertion d'une attitude

À la différence des questions – ouvertes ou fermées – dans lesquelles l'utilisateur est invité à réagir et à se positionner vis-à-vis du contenu propositionnel véhiculé, les assertions permettent, lorsqu'elles réfèrent à une attitude, d'affirmer l'état d'une attitude telle qu'elle est prise en charge par sa source – qu'il s'agisse ou non du locuteur. L'attitude exprimée au sein d'une assertion ne requiert aucunement la réponse de l'interlocuteur pour être actualisée. Néanmoins, face à l'expression d'une telle attitude chez le locuteur, l'interlocuteur aura la possibilité de formuler une réponse de type accord ou désaccord, lui permettant de prendre en charge à son tour en tant que source, l'attitude préalablement exprimée. Considérons l'échange (5), dans lequel l'agent exprime une appréciation dont il est lui-même la source. Quelle que soit la réponse de l'utilisateur, une attitude est bien exprimée et actualisée. En répondant par l'affirmative, en exprimant son accord avec les propos de l'agent, l'utilisateur indique que lui aussi peut être la source de l'appréciation formulée précédemment : il actualise donc ainsi une seconde appréciation.

(5) Agent : "This movie is really bad"  
User : "Yes, indeed"

Néanmoins, ce type d'actualisation semble moins valable pour les expressions de type affect que pour celles de type appréciation ou jugement. Ce type de phénomène est en effet plus approprié à l'expression d'un partage d'une opinion ou d'un avis qu'à celui d'un état émotionnel (voir échange (6)). Du moins, en présence d'une assertion d'affect, la réponse devra être plus précise pour qu'une EASC puisse être exprimée. Ainsi, alors que dans l'échange (6), la réponse de type *yes* semble juste approuver l'état émotionnel exprimé en premier lieu, dans l'échange (7), la réponse plus précise de l'utilisateur permet à celui-ci d'exprimer que son état émotionnel est similaire à celui de l'agent.

(6) Agent : "I'm sad today"  
User : "Yep"  
(7) Agent : "I'm sad today"  
User : "Yep, so do I"

Concernant les valeurs des caractéristiques fondamentales de la seconde attitude, elles sont calculées sur la base de celles de l'attitude première et selon les mêmes calculs que pour les réponses aux questions fermées.

La valeur de la polarité et l'identité de la cible sont conservées en l'état. Si à cette assertion, l'utilisateur répond par la négative alors l'effet sera le même que celui d'une réponse négative à une question fermée, la polarité sera alors inversée.

Assertion	Accord	Désaccord
This movie is good Eval(pos, agent, movieCoref)	"Yes, it is" Eval(pos, user, movie- Coref)	"No, it's not" Eval(neg, user, movie- Coref)
This movie is so bad Eval(neg, agent, movieCoref)	"Yes, it is" Eval(neg, user, movie- Coref)	"No it's not" Appr(neg, user, movie- Coref)
This movie is not really good Eval(neg, agent, movieCoref)	"Yes, it is" Eval(pos, user, movie- Coref)	"No it's not " Eval(neg, user, movie- Coref)

TABLEAU 4.7 – Accord et désaccord portant sur des assertions d'attitude

### 4.3 Formalisation des attitudes strictement collaboratives

Pour une meilleure gestion des EASC pour le système présenté dans le chapitre 6, nous proposons une formalisation des informations sémantiques véhiculées par l'énoncé de l'agent et celui de l'utilisateur.

Question fermée (sans marque de négation) :

$Question(Att(Src = User, Tgt \in T, orientationPol = pos|neg))$

— Réponse de type Accord  $Att(Src = User, Tgt \in T, Pol = orientationPol)$

— Réponse de type Désaccord  $Att(Src = User, Tgt \in T, Pol = inversion(orientationPol))$

Question fermée (avec marque de négation) :

$Question(Att(Src = User, Tgt \in T, orientationPol = pos|neg))$

— Réponse de type Accord  $Att(Src = User, Tgt \in T, Pol = inversion(orientationPol))$

— Réponse de type Désaccord  $Att(Src = User, Tgt \in T, Pol = orientationPol)$

Assertion (sans marque de négation) :

$Assertion(Eval_1(Src_1 = Agent, Tgt \in T, Pol = pos|neg))$

— Réponse de type Accord  $Eval_2(Src_2 = User, Tgt \in T, Pol_2 = Pol_1)$

— Réponse de type Désaccord  $Eval_2(Src_2 = User, Tgt \in T, Pol_2 = inversion(Pol_1))$

Assertion (avec marque de négation) :

$Assertion(Eval_1(Src_1 = Agent, Tgt_1 \in T, Pol_1 = pos|neg))$

— Réponse de type Accord  $Eval_2(Src_2 = User, Tgt_2 = Tgt_1, Pol_2 = inversion(Pol_1))$

— Réponse de type Désaccord  $Eval_2(Src_2 = User, Tgt_2 = Tgt_1, Pol_2 = Pol_1)$

### 4.4 *Small-Talk* vs négociation

L'observation des EASC dans nos deux corpus d'étude n'a pas montré de différences fondamentales entre les deux types d'interaction. Ces expressions sont présentes dans les deux corpus et leurs mécanismes sémantiques suivent les mêmes principes que ceux décrits dans les sections précédentes. La seule différence notable ne concerne pas leur fonctionnement mais leur organisation : dans le corpus de négociation, contrairement au corpus Semaine, l'utilisateur est autorisé à poser des questions, nous pouvons donc y trouver des EASC dans lesquelles l'utilisateur formulera la question ou l'assertion et l'agent la réponse. Cela ne concerne pas notre modèle de détection qui ne s'intéresse qu'aux expressions d'attitude dont l'utilisateur est la source.

En guise d'exemple, nous mettons en évidence ci-dessous les EASC rencontrées dans deux sessions, l'une issue du corpus Semaine, l'autre issue du corpus de négociation. Les énoncés agent et utilisateur permettant l'expression d'une EASC apparaissent de couleur bleue et sont accompagnés dans la marge d'une description sémantique.

O : Hi my name is Poppy	
U : Hello	
O : I'm the eternal optimist	
U : Well you'd need to be because i've had an awful week	
O : <b>An awful week?</b>	<i>Question(Affect(User, Week, Neg))</i>
U : <b>Terrible</b>	<i>Accord : (Affect(User, Week, Neg))</i>
O : That's not good, that's not good to hear em well things will normally get better, they normally eh look up <b>can you think of something that might make you happy</b>	<i>Question(Affect(User, Something, Pos))</i>
U : <b>No</b> , at the minute it's pretty hard but i could always think about holidays I'd like to go away for a holiday in easter but I'm only in the month of january so no point thinking about that	<i>Accord : (Affect(User, Something, Neg))</i>
O : Aha. Where do you think you'll go can you tell me about your holiday?	
U : I'd like to go to San Torrini	
O : Aha. Can you tell me about that?	
U : Ah it's a lovely island in a lovely place but I suspect my husband's arranged to do some work and I'll not be able to get there so that doesn't really help me	
O : Aha but if you think about it, <b>it makes you happy?</b>	<i>Question(Affect(User, SanTorrini, Pos))</i>
U : <b>Yes</b> but I might just end up going to donegal instead or galway or somewhere	<i>Accord : (Affect(User, SanTorrini, Neg))</i>
O : That's good and does do they make you happy?	
U : I haven't been there for a long time but islands generally make me feel happy they kinda like nice peaceful restful places. But a bit of sun would also make me feel happy	
O : That's good	
U : I don't like this weather	
O : Ok and eh so <b>the sun you so you find that uplifting</b>	<i>Question(Affect(User, Sun, Pos))</i>
U : Oh <b>yes</b> , very	<i>Accord : (Affect(User, Sun, Neg))</i>

FIGURE 4.2 – Expressions d'attitude strictement collaboratives. Extrait du corpus Semaine

A : So, what are you most interested in?	
U : uh I'm pretty much equally interested in all these items	
A : Oh really?	
U : They're they're all of uh equal value to me	
A : huh I think you must have a preference for one thing more than the others	
U : I probably prefer the picture	
A : Ok. <b>Do you like the painting?</b>	<i>Question(Affect(User, Painting, Pos))</i>
U : <b>Yes</b>	<i>Accord : (Affect(User, Painting, Neg))</i>
A : <b>Did you want any of the records?</b>	<i>Question(Affect(User, Record, Pos))</i>
U : <b>yeah</b>	<i>Accord : (Affect(User, Record, Neg))</i>
A : Alright well what do you propose?	
U : um I take one more of these and you can have both lamps	
A : Well, that's an interesting idea. let's move the items to show what we're talking about	
U : there you go	

FIGURE 4.3 – Expressions d'attitude strictement collaboratives. Extrait du corpus de négociation

## 4.5 Conclusion du Chapitre 4

Dans ce chapitre, nous nous sommes concentrée sur l'analyse des attitudes au niveau de la paire adjacente. Ce cadre d'analyse nous a permis d'envisager ces expressions comme des actes participatifs, contribuant à la cohésion et à la progression de la conversation. De ce fait il paraît important de pouvoir prendre en compte ce contenu dans le modèle de détection des attitudes de l'utilisateur. Ce point de vue pragmatique sur les attitudes nous a permis d'isoler un type d'attitude : les EASC (expressions d'attitude strictement collaboratives). Ces expressions s'incarnent pleinement dans la paire adjacente et nécessitent une prise en compte conjointe du contenu verbal de l'agent et de celui de l'utilisateur pour être correctement détectées. La particularité de ces expressions est que la valeur de leurs principaux éléments sémantiques est prédéfinie dans le tour de parole de l'agent et instanciée, par la suite, grâce à la réponse de l'utilisateur. En ce qui concerne l'agent, les énoncés participant à ce type d'expressions sont généralement des questions ouvertes ou fermées, ou des assertions dont le contenu propositionnel concerne une attitude dont la source est soit l'utilisateur soit l'agent. En ce qui concerne l'utilisateur, la réponse permet de formuler soit un accord soit un désaccord.

## Chapitre 5

# Les attitudes dans la séquence thématique

### Sommaire

---

5.1	Fondements théoriques à la notion de thème . . . . .	<b>59</b>
5.1.1	De la perspective du contenu à celle du fonctionnement . . . . .	59
5.1.2	Le thème dans les conversations humain-agent . . . . .	59
5.2	Attitudes et thèmes dans des conversations humain-agent . . . . .	<b>60</b>
5.2.1	Séquence thématique : <i>small-talk</i> vs négociation . . . . .	60
5.2.2	Vue d'ensemble des thèmes dans le corpus semaine . . . . .	61
5.2.3	Relations sémantiques entre les cibles et les thèmes du scénario Semaine	62
5.3	Exemple sur une session du corpus Semaine . . . . .	<b>62</b>
5.4	Conclusion du chapitre 5 . . . . .	<b>63</b>

---

DANS les chapitres précédents, nous avons interrogé l'inscription des attitudes dans le tour de parole (Chapitre 3) et dans la paire adjacente (Chapitre 4). L'étude de la structure sous-jacente des énoncés, nous a permis de comprendre des mécanismes syntaxiques et sémantiques de caractérisation de la source, de la cible et de la polarité dans les expressions simples d'attitude (ESA). Afin de dépasser le cadre strict du tour de parole, l'analyse de la paire adjacente, nous a permis d'envisager les attitudes comme des actes participatifs contribuant à la création des connaissances partagées par les locuteurs au cours de la conversation. Nous avons ainsi pu identifier des *expressions d'attitudes strictement collaboratives* (EASC) et analyser leur fonctionnement sémantique. Ce présent chapitre a pour objectif d'ouvrir davantage le cadre d'analyse et d'observer les attitudes dans la séquence thématique courante de la conversation. Puisque les attitudes se présentent comme des actes participatifs ainsi que comme des contributions à la conversation, il est possible de considérer qu'elles entretiennent un lien étroit avec le thème de la conversation. Notre volonté est de déterminer en quoi cette relation peut être exploitée dans le cadre d'une détection des attitudes dans les conversations humain-agent. Dans un premier temps, nous nous concentrons sur la notion de thème telle qu'elle a pu être décrite dans la littérature et telle qu'elle est exploitée dans les conversations humain-agent. Dans un second temps, nous nous intéressons au lien que les attitudes entretiennent avec le thème afin de définir en quoi les informations relatives à ce dernier peuvent améliorer la détection des attitudes.

Publication associée à ce chapitre :

- Langlet, C. and Clavel, C. (2016). Grounding the detection of the user's likes and dislikes on the topic structure of human-agent interactions. *Knowledge-Based Systems*, 106:116–124

## 5.1 Fondements théoriques à la notion de thème

Nous revenons sur les approches les plus courantes adoptées dans la littérature pour définir le thème dans le discours : tant dans son contenu que dans sa capacité à structurer le discours. Dans un second temps, nous évoquons la manière dont le thème est modélisé dans quelques travaux du domaine des agents conversationnels animés.

### 5.1.1 De la perspective du contenu à celle du fonctionnement

D'un point de vue théorique, la notion de thème (*topic* en anglais) a reçu différentes définitions dans la littérature. Comme l'expliquent Brown and Yule (1983), cette notion peut être définie soit dans le cadre d'une analyse de la phrase, soit dans celui de l'analyse du discours. En analyse phrastique, le *thème* se distingue alors du *rhème* (*comment* en anglais) : "the speaker announces a topic and then says something about it" (Hockett, 1958). En analyse du discours, où elle est envisagée dans le cadre d'analyses tant textuelles que conversationnelles, le thème peut être défini comme dont les locuteurs dialoguent ou écrivent (Goutsos, 1997). Il est alors décrit comme un élément discret, comme une unité discursive présente et détectable à la surface d'un énoncé ou du discours (Dittmar, 1992).

À cette perspective du contenu s'ajoute celle du fonctionnement envisageant le thème comme un cadre structurant et participant à l'unité du texte ou de la conversation. Dans cette logique, Downing (2000) interroge la manière dont les séquences thématiques se distinguent les unes des autres tout au long de la progression du discours ainsi que les moyens dont les locuteurs disposent pour introduire de nouveaux thèmes. Dans les conversations en face à face, différents thèmes peuvent être introduits successivement : quand les locuteurs se taisent indiquant ainsi que le thème courant est en cours d'épuisement, un nouveau thème est introduit. Deux stratégies d'introduction de thèmes sont identifiées : une stratégie d'information et une stratégie d'incitation. Le rôle attribué au locuteur varie selon la stratégie adoptée : la stratégie d'information le place dans le rôle de celui qui fournit ou contrôle le nouveau thème, tandis que, la stratégie d'incitation attribue ce rôle à son interlocuteur.

- **Stratégie d'information** : elle consiste à fournir un contenu informatif nouveau considéré comme digne d'intérêt. Cela est généralement suivi par une marque d'attention, verbale ou non-verbale, produite par le second locuteur : *Mm, Did you? ou Really?*. Le thème est ensuite consolidé dans une troisième étape par l'ajout de nouveau contenu
- **Stratégie d'incitation** : le locuteur propose un thème en formulant une question. Cette incitation invite l'interlocuteur à développer le thème par lui-même en respectant le cadre posé par la question.

Dans leur étude, Button and Casey (1985) montrent le même intérêt pour l'introduction des thèmes dans la conversation. Ils mettent en lumière deux types d'introduction qui se rapprochent de ceux mis en valeur par Downing (2000) : *l'incitation* et *la proposition*.

- **Proposition** : le locuteur sélectionne un thème au sein d'un ensemble de thèmes possibles qui paraissent dignes d'intérêt. Le thème n'est pas proposé mais imposé.
- **Incitation** : le locuteur pose un cadre qui incite l'interlocuteur à sélectionner un thème au sein duquel il peut être développé.

La cohérence de la conversation repose donc non seulement sur le contenu, sur ce que disent les locuteurs, mais également sur la manière dont ils ordonnent ce contenu (Schegloff, 1990). La séquentialité et le positionnement des thèmes sont donc des éléments fondamentaux pour l'organisation et la cohérence discursive.

### 5.1.2 Le thème dans les conversations humain-agent

À la différence des conversations humain-humain, dans lesquelles les thèmes sont sélectionnés de manière spontanée, les conversations impliquant un agent conversationnel suivent

un scénario encadrant le déroulement et le propos de la conversation. De plus, la notion de thème étant très fortement corrélée à celle de la cohérence discursive, elle apparaît comme une problématique essentielle dans le cadre du développement de stratégies de dialogue pour les interactions humain-agent.

Afin de préserver la cohérence de conversation, de nombreuses études s'intéressant au développement de stratégies de dialogue interrogent la notion de thème. Deux tendances peuvent être repérées : des travaux s'attachant essentiellement à préserver la cohérence de la conversation ; d'autres fondant la sélection des thèmes sur les modèles utilisateur produits par les modules de détection. Dans cette tendance, Song et al. (2009) proposent un système d'extraction des thèmes dans des pages web à partir des goûts de l'utilisateur. De même, Adam et al. (2010) utilisent les données de l'engagement de l'utilisateur pour décider du moment idéal pour changer de thème. Plus récemment, Glas and Pelachaud (2015a) décrivent un modèle de sélection de thème prenant en compte la perception que l'agent a de l'engagement de l'utilisateur, pour maximiser l'engagement combiné de l'agent et de l'utilisateur.

Dans ce contexte, il apparaît que cohérence discursive, thème et préférences/sentiments sont fortement liés. Si, en ce qui concerne l'agent, nous notons une volonté de développer un moteur dialogue fondant la sélection des thèmes sur la base des goûts de l'utilisateur, en ce qui concerne ce dernier, nous pouvons considérer que les attitudes exprimées respectent le cadre thématique. Rappelons que nous considérons les attitudes comme des actes participatifs collaborant à la progression de la conversation. De fait, mises à part les attitudes autonomes évoquées chapitre 4 et que nous ne traitons pas, les attitudes exprimées auront nécessairement pour cible soit le thème soit l'un de ses sous-aspects. Dans cette logique, afin de faciliter l'identification des cibles, le modèle de détection a la possibilité d'exploiter les informations détenues par le moteur de dialogue sur le thème.

## 5.2 Attitudes et thèmes dans des conversations humain-agent

Afin d'analyser les relations existant entre les cibles et le thème courant, nous observons dans un premier temps la nature des thèmes de notre corpus, ainsi que la manière dont ils sont introduits, puis dans un second temps, nous proposons une description de ces relations.

### 5.2.1 Séquence thématique : *small-talk* vs négociation

Dans les conversations humain-agent, l'agent suit un scénario d'interaction. Le dialogue prend place dans un domaine donné et sa structure est prédéfinie. Selon le type de scénario ainsi que les objectifs de l'interaction, les rôles qu'agent et utilisateur jouent dans la sélection des thèmes varient.

Dans le scénario Semaine, l'interaction a pour objet la discussion informelle devant modifier l'état émotionnel de l'utilisateur. Pour la partie étudiée – opérateur humain vs utilisateur humain – le contenu verbal de l'opérateur est contraint par un script. Il est le seul autorisé à poser des questions : agent et utilisateur sont placés dans un rapport de complémentarité où le premier occupe une position haute et le second une position basse (voir Chapitre 1). C'est l'agent qui dirige la conversation. Ainsi, en utilisant des phrases types à des moments clés de la conversation, il peut gérer les transitions thématiques. Les thèmes peuvent être introduits selon les deux stratégies décrites précédemment (Downing, 2000) : l'opérateur introduit des nouveaux thèmes soit en formulant une affirmation soit en interrogeant l'utilisateur. Avec le second type d'ouverture, il choisit "a possible topic, [...] something tellable or newsworthy by the means of the question. Such an elicitation counts as an invitation to [the user] to develop the topic" Downing (2000). L'opérateur offre ainsi un cadre thématique au sein duquel l'utilisateur pourra définir un thème précis. En revanche, lorsqu'il formule une affirmation en guise d'ouverture thématique, l'utilisateur dispose de moins de liberté.

Dans le corpus de négociation, l'objectif d'interaction est différent : la conversation doit mener à un accord entre l'utilisateur et l'agent sur le partage d'un ensemble d'objets. Les thèmes abordés par l'utilisateur et l'agent ne concernent que les objets de la négociation et correspondent à des phases de négociation sur ces mêmes objets. Dans la mesure où il s'agit d'une interaction fondée sur la compétitivité (voir Chapitre 1), l'utilisateur est à même de poser des questions à l'agent et l'introduire de nouveaux thèmes. Ceci est même un impératif, puisqu'il doit connaître les préférences de l'agent pour être en mesure d'élaborer des stratégies de négociation. Cette dimension compétitive de l'interaction implique également que les séquences thématiques sont beaucoup plus courtes. L'objectif étant de d'obtenir un consensus, les phases de négociation s'enchaînent rapidement et les interlocuteurs ne s'éternisent pas sur un objet ou sur l'un de ses aspects. Il est donc rare de trouver des séquences qui s'étendent sur plusieurs paires adjacentes. Pour cette raison, le travail présenté ci-dessous ne prend en compte que le corpus Semaine. La faible étendue des séquences du corpus de négociation ainsi que la possibilité pour l'utilisateur d'introduire de nouveaux thèmes rendaient en effet trop complexe la modélisation du lien entre attitudes et thème.

### 5.2.2 Vue d'ensemble des thèmes dans le corpus semaine

**Introduction des thèmes :** Dans les sessions observées, les stratégies d'introduction de thèmes par l'agent correspondent aux stratégies décrites ci-dessus : l'opérateur introduit une nouvelle séquence en imposant un thème (stratégie d'information), soit en imposant un cadre thématique au sein duquel l'utilisateur choisit un thème spécifique (stratégie d'incitation). Avec ce second type d'introduction, l'agent fixe les limites de ce dont il va parler, tandis que l'utilisateur sélectionne, dans la zone définie par l'agent, ce dont il sera effectivement question.

**Nature des thèmes :** Quelque soit le type d'introduction, deux catégories de thème ont été trouvées dans le corpus : ceux véhiculant des informations factuelles et ceux véhiculant des informations relatives aux sentiments de l'utilisateur. Dans l'exemple (i) ci-dessous, l'agent procède à l'ouverture du thème en adoptant une stratégie d'information et en imposant comme sujet de conversation l'état émotionnel. Dans l'exemple (ii), la stratégie d'ouverture reste la même, le thème sélectionné est néanmoins factuel (*work*).

- (i) "I'm a happy person, and you?"
- (ii) "Tell me about your work"

Lorsque l'ouverture se fait grâce à une stratégie d'incitation, l'agent détermine le cadre thématique en posant une question. Dans le corpus, nous avons identifié sept cadres thématiques pouvant être sélectionnés par l'agent : *Loisirs*, *Projet en Général*, *Activité professionnelle*, *Etats émotionnels en général*, *Joie*, *Tristesse*, *Colère*. Notons que ces cadres thématiques peuvent se combiner. Par exemple, l'agent peut interroger l'utilisateur sur l'état émotionnel causé par son travail.

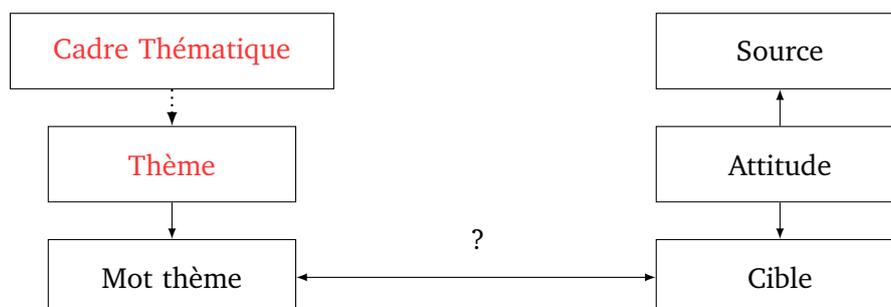


FIGURE 5.1 – Relation sémantique entre la cible et le thème

### 5.2.3 Relations sémantiques entre les cibles et les thèmes du scénario Semaine

Verbalement, un thème s'incarne dans une isotopie sémantique, i.e un ensemble de mots véhiculant des traits sémantiques communs. Cette isotopie se construit dans le discours et n'est pas a priori connue. L'ensemble des mots qui structurent cette isotopie entretiennent avec le thèmes différentes relations lexicales. Nous présentons ici celles qui apparaissent comme elles plus évidentes.

- **synonymie** : la synonymie est la relation sémantique entre mots ayant un sens similaire.
- **hyperonymie/hyponymie** : l'hyperonymie est la relation sémantique hiérarchique d'une unité lexicale à une autre selon laquelle l'extension du premier terme, plus général, englobe l'extension du second, plus spécifique. Par exemple, *félin* est l'hyperonyme de *chat*.
- **holonymie/méronymie** : la méronymie est une relation sémantique entre mots, lorsqu'un terme désigne une partie d'un second terme. Par exemple, *dossier* est le méronyme de *chaise*.
- **proximité sémantique** : des mots sont proches sémantiquement lorsqu'ils peuvent être employés conjointement pour définir une entité, une situation ou un effet. Leur sens n'est pas similaire mais ils participent de la description d'un même univers perceptif.

Dans le cadre de l'analyse des cibles des expressions d'attitude, l'objectif est donc de vérifier dans quelle mesure la cible d'une attitude est un élément de l'isotopie thématique et de définir la nature de cette relation. Pour définir la relation sémantique existant entre une cible et le thème courant, il est nécessaire de disposer du mot ou syntagme caractérisant le thème. À partir de ce mot, il est possible de vérifier les relations sémantiques et lexicales que la cible entretient avec lui. Lorsque l'agent ouvre une séquence thématique en usant d'une stratégie d'information, le mot thème peut être directement obtenu via le moteur de dialogue. En revanche, lorsque l'agent ouvre un cadre thématique, le thème exact choisie par l'utilisateur ne peut pas être récupéré de cette manière. Une analyse lexicale du contenu verbal de l'utilisateur sera nécessaire. Elle permettra, sinon d'obtenir le mot thème, au moins d'identifier son isotopie sémantique. Le chapitre 6 montre l'exemple d'une analyse lexicale de ce type effectué grâce à la ressource *WordNet-Domains*.

## 5.3 Exemple sur une session du corpus Semaine

La figure 5.2 présente une séquence thématique extraite du corpus Semaine. En rouge sont représentés les énoncés permettant d'introduire un cadre ou un thème spécifique : le lieu de vacances pour le premier, et le soleil pour le second. Le premier est un cadre thématique plus restreint que ceux définis plus haut. Nous le considérons comme un cadre car il laisse une marge de manœuvre à l'utilisateur. En bleu est noté le thème sélectionné par l'utilisateur. Nous mettons en gras le membre de l'isotopie sémantique du thème. Les mots *island*, *place* entretiennent une relation d'hyperonymie avec le thème choisi par l'utilisateur *Santorini*. En revanche, les mots **Donegal**, **Galway** entretiennent avec lui une relation de proximité sémantique en ce qu'ils sont tous trois des noms de lieu.

Le second thème introduit appartient à la catégorie *Joie* citée dans la section précédente. Associé à une cible spécifique, *the sun*, ce thème est sélectionné sur la base du contenu verbal antérieur de l'utilisateur (*sun would also make me feel happy*). Nous le considérons malgré tout comme l'introduction d'un nouveau thème : cette relance permet de clôturer la séquence thématique précédente *Lieu de vacances*.

O : Aha. Where do you think you'll go **can you tell me about your holiday?**

Cadre - Lieu de vacances - Loisirs

U : I'd like to go to **San Torrini**

Thème spécifique choisi par l'utilisateur

O : Aha. Can you tell me about that?

U : Ah it's a lovely **island** in a lovely **place** but I suspect my husband's arranged to do some work and I'll not be able to get there so that doesn't really help me

O : Aha but if you think about it, it makes you happy?

U : Yes but I might just end up going to **Donegal** instead or **Galway** or **somewhere**

O : That's good and does do they make you happy?

U : I haven't been there for a long time but **islands** generally make me feel happy they kinda like nice peaceful restful **places**. But a bit of sun would also make me feel happy

O : That's good

U : I don't like this weather

O : Ok and eh so **the sun you so you find that uplifting**

Nouveau thème proposé par l'agent

U : Oh yes, very

FIGURE 5.2 – Structure thématique dans une session du corpus Semaine

## 5.4 Conclusion du chapitre 5

Dans ce chapitre, nous avons élargi le cadre d'étude afin d'observer les attitudes dans le cadre de la séquence thématique. Afin d'améliorer le traitement des cibles des attitudes, nous sommes interrogée sur la manière d'exploiter les informations détenues par le moteur de dialogue concernant le thème courant de la conversation. Le thème dans la littérature a reçu différentes définitions et a été étudié selon différentes approches. Analysé comme une unité de discours, il peut être également défini comme un cadre participant au fonctionnement de la conversation. De ce point de vue, sont alors étudiées les stratégies d'introduction du thème. Cette approche a retenu notre attention, car elle peut donner des pistes sur le rôle que l'agent et l'utilisateur peuvent jouer respectivement pour introduire un thème. Dans les conversations humain-agent, l'introduction du thème tient une place importante. Le choix de celui-ci repose sur des stratégies d'interaction devant guider la conversation. Dans certains travaux, le choix du thème est décidé sur la base des goûts et des préférences de l'utilisateur. Ce type d'approche encourage notre démarche consistant en un ancrage de l'analyse des attitudes de l'utilisateur dans la structure thématique. Afin de permettre cet ancrage, nous avons étudié la structure thématique de l'un de nos deux corpus d'étude, le corpus Semaine (le corpus de négociation nous est apparu comme moins adapté). Dans les conversations de Semaine, l'agent peut ouvrir une nouvelle séquence thématique en utilisant deux stratégies définies dans la littérature : une stratégie d'incitation ou une stratégie d'information. Dans le premier cas, il impose un cadre thématique au sein duquel l'utilisateur choisit un thème, dans le second cas, le thème est imposé. Dans le cadre de la première stratégie, il est nécessaire d'identifier le thème choisi par l'utilisateur afin de pouvoir effectuer une analyse des liens sémantiques existant entre les cibles des attitudes et

le thème courant. Pour cela, une étude de l'isotopie sémantique (ensemble de mots véhiculant des traits sémantiques communs) incarnant verbalement le thème permet d'obtenir des informations intéressantes. Concernant les liens sémantiques, quatre relations ont été identifiées : synonymie, hyperonymie/hyponymie, holonymie/méronymie et proximité sémantique.

## Troisième partie

# Détection des *likes* et *Dislikes* : de la détection à l'évaluation



---

## Synthèse de la partie III

Cette partie est composée de deux chapitres : le chapitre 6 est consacré à la description de notre système de détection des likes et dislikes et le chapitre 7 décrit les trois évaluations réalisées pendant notre thèse. Les expressions de *likes* et *dislikes* traitées dans cette partie regroupent toutes les expressions d'attitude permettant de référer aux goûts de l'utilisateur. Cette catégorie doit être considérée comme contextuelle et orientée vers un objectif applicatif. Le choix de ce type d'expressions a été guidé par l'importance dans le domaine des agents conversationnels animés de la dimension du *liking*.

Le système décrit dans le chapitre 6 s'organise en trois principaux traitements. Chacun d'eux est consacré à l'analyse de l'un des trois cadres d'analyse décrits dans la partie précédente : le tour de parole, la paire adjacente et la séquence thématique. Pour chacun de ces niveaux d'analyse, le système procède à un traitement ascendant (*bottom-up*).

- Au niveau du tour de parole, nous produisons un ensemble de grammaires fournissant une analyse tant syntaxique que logico-sémantique des énoncés et permettant d'identifier des expressions de *likes* et de *dislikes*, de connaître les syntagmes référant à leur cible et à leur source et de calculer leur polarité ;
- Au niveau de la paire adjacente, nous fournissons une chaîne de traitement intégrant différentes règles sémantiques pour l'identification d'expressions de *likes* et *dislikes* strictement collaboratives. La chaîne de traitement proposée combine analyse du tour de parole de l'agent et analyse du tour de parole de l'utilisateur et fournit en sortie l'ensemble des *likes* et *dislikes* strictement collaboratifs trouvés dans la paire adjacente et définit la valeur de leurs attributs sémantiques *polarité*, *cible* et *source*.
- Au niveau de la séquence thématique, nous proposons une architecture d'analyse lexicale permettant de définir la nature des relations sémantiques existant entre les cibles des *likes* et *dislikes* détectés et le thème courant de la conversation.

Dans le chapitre 7, trois évaluations sont présentées. Les deux premières ont été effectuées sur le corpus *Semaine*, la troisième sur le corpus de négociation. L'une des évaluations sur le corpus *Semaine* et celle sur le corpus de négociation ont pour objectif de tester les performances du système dans le cadre du tour de parole et de la paire adjacente. Il s'agit d'évaluations quantitatives ayant nécessité la constitution d'un ensemble d'annotations de référence. Les tâches d'annotation proprement dites ont été réalisées grâce à deux plateformes d'annotation développées par nos soins. Les deux évaluations montrent des résultats intéressants.

- pour l'évaluation *Semaine*, nous obtenons un kappa de 0.61 pour la présence d'un *like* ou d'un *dislike*, et alpha de Cronbach de 0.67 pour la détection du nombre d'expressions présentes dans la paire. Quant au calcul de la polarité pour chaque expression, nous obtenons un kappa de 0.81 ;
- pour l'évaluation *Négociation*, nous obtenons un kappa de 0.62, pour la détection des *likes*, sans distinction sur les objets de la négociation, et un kappa de 0.53, pour les *dislikes*, toujours sans distinction sur les objets de la négociation.



## Chapitre 6

# Détection des *likes* et *dislikes* : du tour de parole à la séquence thématique

### Sommaire

---

6.1	Méthodologie et objectifs applicatifs . . . . .	72
6.1.1	Méthodologie : une approche <i>corpus-driven</i> . . . . .	72
6.1.2	Des attitudes aux <i>likes</i> et <i>dislikes</i> . . . . .	74
6.2	Tour de parole de l'utilisateur : détection des expressions simples de <i>likes</i> et <i>dislikes</i> . . . . .	76
6.2.1	Analyse lexicale . . . . .	77
6.2.2	Analyse syntagmatique . . . . .	79
6.2.3	Analyse phrastique . . . . .	81
6.3	Paire adjacente : détection des <i>likes</i> et <i>dislikes</i> strictement collaboratifs . . . . .	83
6.3.1	Tour de parole de l'agent . . . . .	83
6.3.2	Tour de parole de l'utilisateur . . . . .	84
6.3.3	Calcul des propriétés sémantiques . . . . .	84
6.4	Séquence thématique : optimiser le traitement des cibles . . . . .	86
6.4.1	Niveau A : relations entre les thèmes et les domaines de <i>WordNet-Domains</i> . . . . .	86
6.4.2	Niveau B : détection de mots-thèmes . . . . .	87
6.4.3	Niveau C : vers un modèle utilisateur . . . . .	87
6.4.4	Exemple sur une séquence du corpus <i>Semaine</i> . . . . .	90
6.5	Conclusion du chapitre 6 . . . . .	92

---

DANS les chapitres précédents, nous avons observé et formalisé les expressions d'attitude telles qu'elles se manifestent dans un cadre conversationnel. Dans ce chapitre, nous quittons le domaine de l'analyse théorique pour aborder la question proprement opérationnelle de notre travail, celle de la méthode de détection des attitudes dans un contexte de conversation humain-agent. Nous présentons ici l'architecture de notre système de détection ainsi que l'ensemble des règles que celui-ci intègre dont les performances seront exposées dans le chapitre suivant (chapitre 7). Ce modèle se fonde sur les observations préalablement déployées mais, afin de prendre en compte un certain nombre d'impératifs applicatifs, il se concentre sur un sous-phénomène linguistique celui des *likes* et *dislikes*. Cet ensemble d'expressions peut être considéré comme une sous-catégorie transverse à celle des attitudes. Forte de la modélisation offerte par le modèle de Martin and White (2005) et de la caractérisation qu'il propose des différentes expressions d'attitude, nous avons sélectionné parmi ces expressions celles qui répondaient aux impératifs applicatifs posés par une détection dans le cadre de conversation humain-agent.

L'architecture du modèle s'articule autour de trois principaux traitements, dédiés pour chacun à l'un des trois cadres d'analyse mis en valeur par les précédents chapitres. Le système dans son architecture est conçu comme une chaîne de traitement *bottom-up* analysant, dans un premier temps, le niveau inférieur du tour de parole, puis dans un deuxième temps, celui de la paire adjacente et enfin, dans un troisième temps, celui de la séquence thématique. Cette approche *bottom-up* est également celle mise en place à l'intérieur de ces différents niveaux d'analyse. Pour chacun d'eux, le système procède tout d'abord à l'analyse des éléments minimaux, comme par exemple les unités lexicales, afin de pouvoir *in fine*, analyser les éléments de niveau supérieur, comme par exemple les structures phrastiques ou les mécanismes sémantiques qui s'établissent entre les énoncés. Les traitements présentés dans ce chapitre ont été développés pour les deux types d'interaction déjà précédemment envisagés : le *small-talk* et la négociation. Le processus de traitement respecte les mêmes fondamentaux : une analyse ascendante qui s'intéresse successivement au niveau lexical, au niveau syntagmatique et au niveau phrastique. Néanmoins, les particularités inhérentes à leurs scénarios respectifs imposent quelques différences de traitement que nous aborderons tout au long du chapitre.

Après une présentation de notre méthodologie et des impératifs applicatifs nous ayant conduit à définir la notion de *likes* et de *dislikes* (section 6.1), nous présentons l'ensemble des règles développées pour la détection des *likes* et *dislikes* simple (L&DS) dans le cadre du tour de parole de l'utilisateur. Les L&DS ont un fonctionnement similaire à celui des attitudes simples décrites dans le chapitre 3. Nous abordons ensuite la question des expressions de *likes* et *dislikes* strictement collaboratives (L&DSC) et présentons la méthode développée pour les localiser au sein de la paire adjacente et déterminer la valeur sémantique de leurs attributs *source*, *cible* et *polarité*. Enfin, nous détaillons l'ensemble des traitements réalisés au sein de la séquence thématique pour déterminer les potentielles relations sémantiques entre la cible et le thème courant de la conversation.

---

Publications associées à ce chapitre :

- Langlet, C. and Clavel, C. (2015b). Improving social relationships in face-to-face human-agent interactions: when the agent wants to know user’s likes and dislikes. In *Proceedings of the 53rd Annual Meeting of the Association for Computational Linguistics and the 7th International Joint Conference on Natural Language Processing (Volume 1: Long Papers)*, volume 1, pages 1064–1073
- Langlet, C. and Clavel, C. (2015a). Adapting sentiment analysis to face-to-face human-agent interactions: from the detection to the evaluation issues. In *Affective Computing and Intelligent Interaction (ACII), 2015 International Conference on*, pages 14–20. IEEE
- Langlet, C. and Clavel, C. (2016). Grounding the detection of the user’s likes and dislikes on the topic structure of human-agent interactions. *Knowledge-Based Systems*, 106:116–124
- Langlet, C. and Clavel, C. (2018). Detecting user’s likes and dislikes for a virtual negotiating agent. In *20th ACM International Conference on Multimodal Interaction (ICMI 2018)*, Boulder, Colorado, (Submit)

## 6.1 Méthodologie et objectifs applicatifs

Le système décrit dans ce chapitre se compose d'un ensemble de règles formalisant des caractéristiques sémantiques, syntaxiques et pragmatiques rattachées aux expressions d'attitude et mises en évidence dans la partie précédente. Ces règles ainsi que la chaîne qui les ordonne et les articule, constituent le cœur de notre système. Entre le développement de ce système et les observations présentées dans la partie précédente (chapitre 3, chapitre 4 et chapitre 5), un travail d'ordre méthodologique et une caractérisation des objectifs applicatifs ont été nécessaires. Dans un premier temps, cette section présente un schéma d'annotation des attitudes en conversation, qui nous a permis de faire converger les observations précédemment menées vers un modèle computationnel, et dans un second temps, elle présente la manière dont, au sein des expressions d'attitude, nous avons sélectionné celles qui nous paraissaient les plus adaptées aux contextes d'application visés.

### 6.1.1 Méthodologie : une approche *corpus-driven*

Sur la base des observations menées dans les chapitres précédents (chapitre 3, chapitre 4 et chapitre 5), nous avons développé un schéma d'annotation des attitudes en contexte conversationnel (voir la figure 6.1). Ce schéma a pour objectif de vérifier de manière systématique et sur un ensemble de sessions, les caractéristiques spécifiques au contexte conversationnel. Cette phase d'annotation et le schéma sur lequel elle se fonde nous ont permis d'aboutir au modèle computationnel présenté dans la suite du chapitre. Le schéma d'annotation utilisé regroupe les différentes caractéristiques des expressions d'attitude et prend en compte leur contexte d'expression (tour de parole, paire adjacente et séquence thématique).

Au niveau du tour de parole de l'utilisateur, le schéma intègre les caractéristiques sémantiques propres aux expressions d'attitude. Des unités d'annotation sont ainsi réservées au syntagme conférant à l'énoncé son caractère axiologique ou affectif, au syntagme référant à la source et à celui référant à la cible. Des traits sémantiques sont également utilisés pour spécifier la polarité et la sous-catégorie d'attitude. Pour cette dernière, la catégorie *affect* est conservée comme telle, tandis que les catégories *appréciation* et *judgement* ont été fusionnées au sein d'une même catégorie *évaluation*.

Afin de mettre en relief le caractère conversationnel des expressions d'attitude étudiées et la manière dont celles-ci peuvent être déterminées par le contenu verbal de l'agent, les énoncés de ce dernier sont annotés au regard des actes illocutoires qu'ils produisent et des attitudes qu'ils expriment. Les actes illocutoires sont annotés selon la classification de Searle (Searle, 1976) qui comprend cinq catégories : les actes représentatifs (*It's raining*) ; les actes directifs (*I order you to leave, Why are you living ?*) ; les actes commissifs (*I promise to pay you the money*) ; les actes expressifs (*I apologize for stepping on your toe*) ; les actes de déclaration (*You're fired*). Quant aux attitudes exprimées par l'agent, le schéma d'annotation leur consacre des éléments spécifiques : unité d'annotation pour la cible et la source, traits sémantiques pour la sous-catégorie d'attitude et pour la polarité. Pour vérifier la proximité sémantique entre les attitudes de l'utilisateur et celles de l'agent, nous avons créé un lien permettant de relier la cible de l'utilisateur à celle de l'agent, lorsque toutes deux réfèrent à la même entité. De plus, à un plus haut niveau de granularité, chaque attitude exprimée par l'utilisateur est associée, par une relation sémantique dédiée, à l'énoncé de l'agent qui la précède.

Enfin, une unité est consacrée à l'annotation du thème courant. Nous considérons le thème comme une séquence, qui peut se répartir sur plusieurs tours de parole. Un même *thème* peut inclure différents *sous-thèmes*, référant à ses sous-aspects. Le modèle d'annotation intègre deux unités d'annotation dédiées aux *thèmes* et aux *sous-thèmes*. Malgré la nature séquentielle du *thème*, il est ici pris en compte de manière atomique : n'est annotée que la première occurrence du mot référant au thème ou au sous-thème. Chaque sous-thème est relié à son thème de référence via une relation dédiée. Un trait est ajouté aux thèmes et aux sous-thèmes pour indiquer

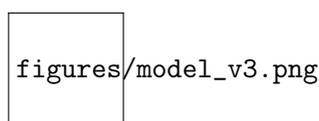


FIGURE 6.1 – Schéma d’annotation

si ceux-ci ont été introduits par l’agent ou l’utilisateur. Lorsqu’une cible réfère à un thème ou sous-thème, elle est alors reliée à celui-ci par une relation spécifique.

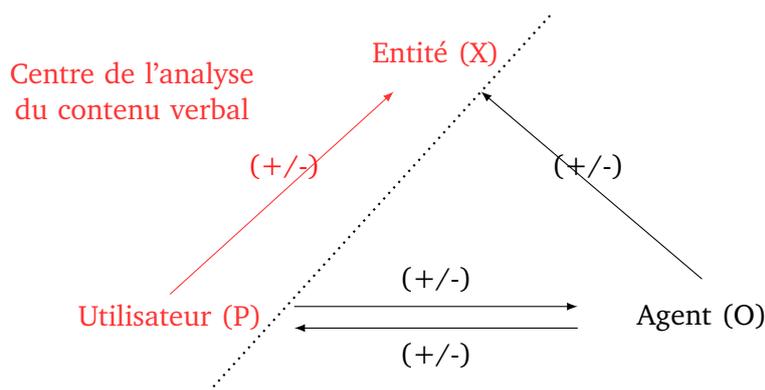
En guise de résumé, la figure 6.1 présente la manière dont deux tours de parole sont annotés. L’énoncé de l’agent a une forme interrogative, il est donc annoté comme *acte directif*. Le contenu propositionnel de la question réfère à une évaluation exprimée par le verbe “like”, dont la source est l’utilisateur et la cible “outdoors activities”. Quant à la phrase de l’utilisateur, le syntagme “don’t like” y est annoté comme *indice d’attitude*, “I” comme *source* et “being indoors” comme *cible*. Afin de mettre en valeur le lien ontologique existant entre elles, une relation est posée entre les deux cibles.

Une première annotation a été effectuée sur seize sessions du corpus Semaine. Un seul annotateur a procédé à l’annotation. Comme outil d’annotation, nous utilisons la plateforme *Glozz* car celle-ci permet d’intégrer des unités (« séquences d’éléments textuels adjacents »), des relations (« rapports binaires entre deux unités ») et schémas (« configurations textuelles complexes récurrentes impliquants unités et relations »).

Le nombre de tour de parole varie selon les sessions (132 pour la plus longue, 38 pour la plus courte). Parmi les 450 énoncés de l’agent annotés, 46% expriment un acte directif, 42% un acte expressif et 12% un acte représentatif. Aucun acte déclaratif ou commissif n’a été trouvé. Parmi l’ensemble des énoncés de l’agent, 339 réfèrent à une attitude (187 à un affect, 152 à une évaluation). Du côté de l’utilisateur, 238 attitudes ont été annotées : 44% sont notées comme affect, 56% comme évaluation (jugements et appréciations).

Au regard de l’annotation manuelle effectuée sur le corpus Semaine, il est possible d’avoir un premier aperçu de la manière dont l’agent peut influencer les expressions d’attitude de l’utilisateur, et comment ce dernier se place dans une relation de collaboration dialogique avec l’agent. Comme nous avons déjà expliqué (Langlet and Clavel, 2014b), sur un plan pragmatique, la nature de l’acte illocutoire précédant les expressions d’attitude de l’utilisateur est un premier indicateur. En effet, si les actes directifs et expressifs de l’agent ont chacun un nombre significatif d’occurrences dans l’ensemble du corpus annoté, au regard de la relation qui les associe aux attitudes exprimées par l’utilisateur, les actes directifs prévalent. Il apparaît d’ailleurs qu’une majorité de ces actes directifs sont des questions. Ainsi, en tant que requêtes, ces énoncés peuvent facilement produire une réaction affective de la part de l’utilisateur. De plus, parmi ces questions, certaines concernent directement des attitudes de l’utilisateur. Ensuite, sur le plan du contenu propositionnel, il est à noter que, lorsqu’une attitude de l’utilisateur est précédée par un énoncé de l’agent contenant lui aussi une expression d’attitude, les polarités concordent généralement. De plus, concernant les cibles, un quart des cibles évaluées par l’utilisateur réfère à une cible présentes dans l’énoncé de l’agent.

Enfin, concernant les thèmes, parmi les 61 unités lexicales annotées comme thème et les 51 annotées comme sous-thème. 57% de thèmes sont introduits par l’agent, 30% par l’utilisateur et 13% par une collaboration entre l’agent et l’utilisateur. Par exemple, lorsque l’agent demande à l’utilisateur “Where is the best wake up you ever had?”, et que l’utilisateur répond “in a tent in the Kilimanjaro”, la question de l’agent ouvre un cadre thématique, au sein duquel la réponse de l’utilisateur définit un thème spécifique. Concernant, le lien entre attitudes et thèmes, parmi les cibles évaluées par l’utilisateur, 32% réfèrent à un thème ou sous-thème introduit par l’agent, 40% à thème ou sous-thème introduit par un utilisateur et 28% à un thème ou sous-thème issu d’une collaboration entre les interlocuteurs.

FIGURE 6.2 – *Balance Theory* de Heider

### 6.1.2 Des attitudes aux *likes* et *dislikes*

Dans le domaine des agents conversationnels animés, la modélisation des relations interpersonnelles se fonde sur l'utilisation de différentes dimensions, dont l'une des plus souvent citées est celle du *liking*. Sa définition repose sur la *Balance Theory* de Heider, qui formalise la façon dont les relations entre individus sont expérimentées cognitivement lorsqu'elles impliquent une entité impersonnelle (Heider, 1958). Le modèle prend en compte une personne P, au centre de l'analyse, une seconde personne O et une entité X, qui peut être un événement, un processus ou un objet physique (voir la figure 6.2). Tous trois entretiennent des relations représentées par un triangle dont les sommets représentent P, O et X. Les côtés sont les relations de *liking* existant entre eux : le *liking* de P envers O, celui de P envers X et celui de O envers X. Pour Heider, un état équilibré existe si les trois relations sont positives à tous égards ou si deux sont négatives et une positive Heider (1958). Ainsi, si P aime X, O aime X et P aime O, alors l'état entre P, O et X est équilibré. De même, si P et O n'aiment pas X et P aime O, alors l'état est équilibré. Dans le cas contraire, l'énoncé est déséquilibré, cela peut être une source de tension pour P. Cette théorie suppose que pour réduire la tension, les personnes impliquées dans ce type de relation tendent à garder des états équilibrés. Certains scénarios d'interaction se fondent sur l'hypothèse que la relation de *liking* de l'utilisateur (P) envers l'agent (O) est déterminée par le sentiment que l'agent et l'utilisateur ont vis-à-vis d'une entité impersonnelle (X). Si l'agent et l'utilisateur partagent le même sentiment envers l'entité impersonnelle (X), alors la relation de *liking* de l'utilisateur envers l'agent devrait être positive.

Pour une première version de notre système, nous avons donc choisi de nous concentrer sur la détection des expressions d'attitude informant sur les goûts de l'utilisateur et utiles à la détermination de la valeur de la relation de *liking* envers l'agent. Nous avons regroupé ces différentes expressions au sein d'une même catégorie : les *likes* et *dislikes* (L&D). Cette catégorie doit être considérée comme une catégorie contextuelle et applicative permettant de circonscrire les expressions à détecter au regard des besoins spécifiques de l'agent, et non comme une catégorie conceptuelle et théorique utile à la modélisation d'un phénomène linguistique. À partir du modèle de Martin and White (2005), nous avons sélectionné les expressions permettant au locuteur, dans notre cas l'utilisateur, d'indiquer ce qu'il aime, n'aime pas, ce qui l'intéresse ou l'indiffère, ce qui le subjugué et ce qui le répugne. Ces expressions sont issues des trois classes définies par le modèle : affects, jugements et appréciations. Prenons l'exemple des phrases ci-dessous :

- (i) "This movie makes me sad"
- (ii) "This painting is a masterpiece"

Tandis que sur le plan théorique, la première est considérée comme un affect et la seconde comme une appréciation, sur le plan applicatif toutes deux peuvent être envisagées comme des indices de L&D. Cependant, parmi les expressions d'attitudes au sein desquelles notamment

aucune cible ou déclencheur ne sont spécifiés, ne peuvent être envisagées comme des L&D. Par exemple, “*I’m very happy*” fait uniquement référence à un état affectif et ne donne aucune indication sur les goûts de l'utilisateur. Ce type d'énoncé est alors exclu de la catégories des *likes* et *dislikes*.

**Likes et Dislikes dans le Cadre de la Négociation** Néanmoins, si cette catégorie est valable dans le cadre de différents scénarios, il est à noter que certains énoncés peuvent s'y ajouter dans certains contextes spécifiques. Dans le cadre d'un scénario de négociation, un ensemble plus large d'énoncés peut être considéré comme des indices de *likes* et *dislikes*. Chaque locuteur impliqué dans une négociation a un objectif de communication spécifique : obtenir les objets qui l'intéressent. Le dialogue doit être efficace et doit conduire à cet objectif. Dans un tel contexte, les expressions de *likes* et *dislikes* peuvent avoir un impact sur le résultat final de la négociation. Dans ce cadre spécifique, nous considérons ces expressions, non comme reflétant une composante de la personnalité de l'utilisateur, mais comme un outil permettant de tirer parti de la négociation. Le locuteur exprime des goûts ou des aversions pour servir un objectif pragmatique.

Cette nature particulière est illustrée par le type de phrases pouvant donner un indice sur les goûts de l'utilisateur. Deux types d'expression peuvent référer à des *likes* et *dislikes* : d'abord, des expressions d'attitude telles que décrites par Martin and White (2005), ensuite, des phrases liées à une déclaration volitive. Tandis que le premier type d'expressions se réfère aux propriétés de l'objet et aux sentiments que celui-ci procure chez l'utilisateur, le second permet d'exprimer *les souhaits et volonté de l'utilisateur*. Dans la phrase, “*I want the records*”, l'utilisateur n'attribue pas de valeur axiologique ou affective à l'objet, il explique simplement sa volonté concernant le résultat de la négociation. Nous considérons que ces deux types de phrases sont sémantiquement corrélés : en disant qu'elle aime un objet, la personne qui négocie suggère également son souhait de l'obtenir ; tandis qu'en disant qu'elle désire acquérir un objet spécifique, elle suggère qu'elle l'aime ou lui donne une valeur positive. Ce type d'énoncés a donc été considéré par notre système de détection.

## 6.2 Tour de parole de l'utilisateur : détection des expressions simples de *likes* et *dislikes*

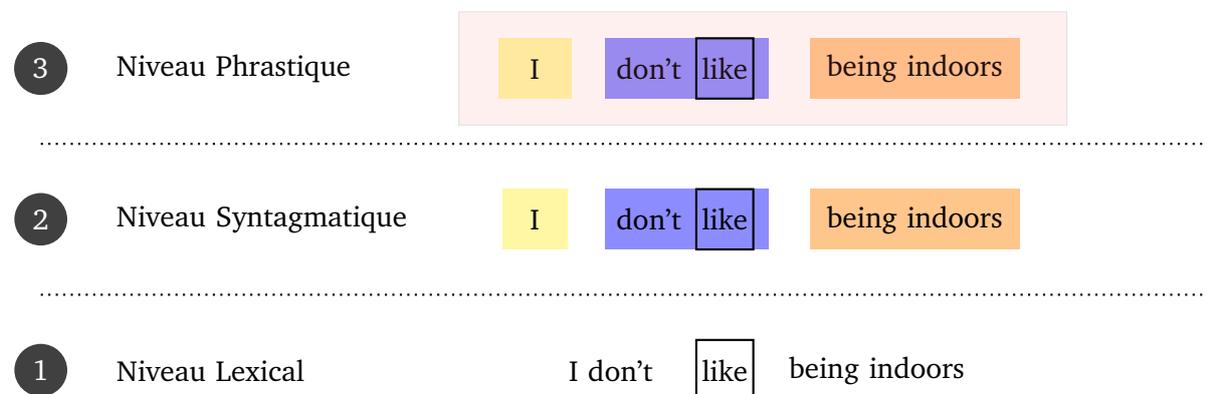


FIGURE 6.3 – Traitement des expressions simples de *likes* et *dislikes* dans le tour de parole de l'utilisateur

Les traitements présentés dans cette section ont comme contexte d'analyse le tour de parole de l'utilisateur. Pour chaque tour de parole, l'objectif est donc de fournir la liste des *likes* et *dislikes* qui y sont exprimés. L'objectif de l'analyse est d'explorer la structure logico-sémantique des énoncés en se fondant sur le principe de compositionnalité sémantique, c'est-à-dire le principe qui définit le sens des énoncés comme constitué du sens des différents éléments qui les composent et des relations hiérarchiques qui les structurent. Afin d'intégrer ce principe, notre chaîne de traitement s'organise comme un processus d'analyse ascendant : après une phase de pré-traitements (*tokenisation* et *POS-tagging*), le système procède à une analyse lexicale, puis à une analyse syntagmatique et enfin à une analyse phrastique. Les sorties produites par chaque niveau sont ensuite exploitées par les niveaux supérieurs. Chacune de ces phases d'analyse exploite des ressources spécifiques : lexicale et grammaticales locales, implémentées sous formes de transducteurs dans la plateforme *Unitex*<sup>1</sup>. Les traitements appliqués au tour de parole reposent sur le même principe de fonctionnement quelque soit le type d'interaction (négociation ou *small-talk*). Néanmoins, certains traitements sont spécifiques au contexte de la négociation. Pour simplifier la lecture, nous présentons chaque étape de l'analyse dans son fonctionnement général et précisons, pour chacune d'elles, le contexte d'analyse.

1. Le détail des grammaires peut être trouvé en Annexe F et leur implémentation sous *Unitex* est expliquée en Annexe E

### 6.2.1 Analyse lexicale

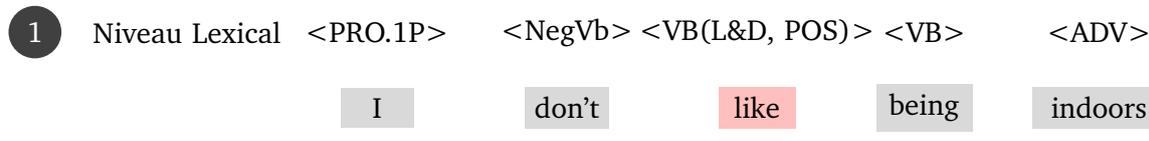


FIGURE 6.4 – Analyse lexicale du tour de parole

L'analyse lexicale s'intéresse à la détection de lexies clés pour la reconnaissance d'expressions de *like* ou *dislike*. Elle consiste en une projection, sur l'énoncé de l'utilisateur, de lexiques fournissant un jeu d'annotations aux différentes unités reconnues. Selon le phénomène ciblé par le lexique, les annotations correspondent à des informations sémantiques ou grammaticales. Trois analyses lexicales différentes sont effectuées à ce niveau : l'une est commune aux applications *small-talk* et négociation, les deux autres sont spécifiques au contexte de la négociation. La première permet de détecter des lexies référant à un affect, une appréciation ou un jugement ; la seconde, les objets de la négociation et donc les potentielles cibles, et la troisième des verbes permettant d'exprimer une volition.

#### 6.2.1.1 Utilisation de *WordNet-Affect* (*small-talk* et négociation)

Pour identifier les lexies permettant d'exprimer un *like* ou un *dislike*, nous avons choisi d'exploiter la ressource *WordNet-Affect* (Valitutti, 2004). Parmi l'ensemble des ressources lexicales existantes (comme par exemple, Baccianella et al. (2010) ou Pennebaker et al. (2001)), notre choix a été en partie déterminé par la nature de son recouvrement lexical, proche de la notion de *likes* et *dislikes* telle que présentée à la section 6.1.2. Néanmoins, afin d'améliorer sa précision, nous ne l'avons pas utilisée dans sa globalité. *WordNet-Affect* se présente comme une extension de *WordNet Domains*. La ressource consiste en une hiérarchie de domaines affectifs (*affect domains* en anglais), indépendants de la hiérarchie initiale de *WordNet-Domains*. Chaque domaine affectif recense plusieurs synsets, issus de *WordNet*, référant à la dimension affective représentée par le domaine. Au sommet de cette hiérarchie, qui s'inspire en partie du travail de Clore et al. (1987), on compte six domaines affectifs principaux : *mental-state*, *physical-state*, *behaviour*, *situation*, *signal*, *trait*, *sensation*. Le domaine des *mental-states*<sup>2</sup> comporte trois sous-domaines : *affective-states*, *cognitive-states* et *cognitive-affective-states*. Enfin, le domaine des *affective-states* se compose de deux sous-domaines : *mood* et *emotion*. Parmi l'ensemble des synsets, n'ont été choisis que ceux dont les lexies pouvaient être utilisées pour l'expression d'un *like/dislike* et appartenant au domaine *emotion* ou à deux de ses quatre sous-domaines : *positive-emotion* et *negative-emotion*.

Les catégories utilisées dans *WordNet-Affect* ne correspondent pas aux catégories développées par Martin and White (2005). Les catégories *positive-emotion*, *negative-emotion* intègrent des unités lexicales référant à des émotions ou affects (par exemple, *sad*, *happy*, *scared*), mais intègrent également des lexies pouvant être employées dans des énoncés exprimant un jugement ou une appréciation (par exemple *beautiful*, *brave*, *nice*, *valuable*). Il est également important de préciser à ce stade que les unités lexicales ne peuvent pas à elles seules caractériser une expression d'attitude. Le contexte au sein duquel elles s'intègrent est également déterminant. Ainsi la polarité, la cible et la source ne sont définies qu'à un niveau phrastique. Les deux exemples ci-dessous appartiennent à deux catégories différentes : jugement pour le premier, appréciation pour le second. Néanmoins, ils emploient le même adjectif.

2. Nous ne reportons pas ici l'ensemble de la hiérarchie juste la branche pertinente pour notre travail. La hiérarchie dans sa globalité se trouve en annexes du mémoire

- (i) He is a bad man
- (ii) It is a bad movie

### 6.2.1.2 Indices lexicaux spécifiques au contexte de la négociation

**Indices de volition :** La détection dans un contexte de négociation implique également de pouvoir détecter des éléments lexicaux constitutifs de L&DS de type volitif. Ces indices lexicaux sont essentiellement verbaux (*want, wish, get, have*). Là encore, nous avons constitué manuellement un court lexique sur la base des observations que nous avons pu menées sur notre corpus d'étude.

**Objets de la négociation :** Dans le cadre d'une application à un corpus de négociation, le système a été conçu pour pouvoir détecter des Lconcernant uniquement les objets de la négociation. Au niveau lexical, l'objectif est donc de détecter tous les noms (simples ou composés) pouvant désigner les objets impliqués dans la négociation entre l'agent et l'utilisateur, alors que pour le *small-talk* les cibles seront traitées de différentes façons (voir la section 6.2.3). Pour une première expérimentation, nous avons conçu manuellement un court lexique de synonymes. Nous nous sommes fondée sur la liste des objets du corpus d'évaluation (DeVault et al., 2015).

Valeur Sémantique	Cat. grammaticale	Représent. formelle	Exemple	Nb. entrées	
L&D	Adjectifs	<i>Adj(L&amp;D)</i>	<i>sad, nice</i>	980	Small-talk/négociation
	Noms	<i>N(L&amp;D)</i>	<i>joy</i>	563	
	Verbes	<i>Vb(L&amp;D)</i>	<i>like, hate</i>	434	
Obj. négo	Nom	<i>N(ObjNego)</i>	<i>painting, lamp, artwork</i>	24	Négociation
Volition	Verbe	<i>Vb(Vol)</i>	<i>want, get</i>	12	

TABLEAU 6.1 – Synthèse des indices lexicaux annotés

Le tableau ci-dessus représente les traitements lexicaux appliqués au tour de parole. Chacune des trois lignes principales concerne un traitement spécifique. Pour chacun nous indiquons les catégories grammaticales des unités pouvant être annotées, la représentation formelle du label produit et le nombre d'entrées des lexiques utilisés.

### 6.2.2 Analyse syntagmatique

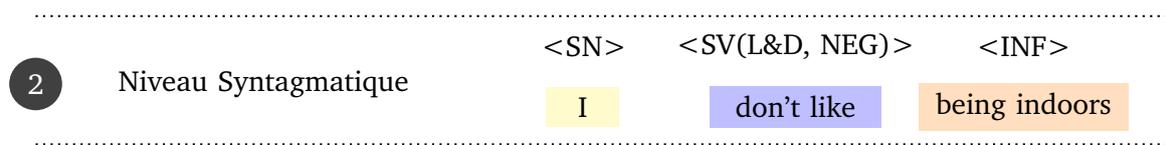


FIGURE 6.5 – Analyse syntagmatique dans le tour de parole de l'utilisateur

L'objectif de l'analyse syntagmatique est d'identifier pour chaque unité lexicale, le syntagme au sein duquel elle s'intègre. Les syntagmes sont envisagés comme des blocs qui, potentiellement, peuvent composer les expressions de L&D, telles que celles-ci seront identifiées au cours de l'analyse phrastique. L'objectif ici est donc de définir les limites syntaxiques de ces blocs, de leur attribuer une catégorie grammaticale (syntagme nominal, adjectival, verbal, etc), de leur attribuer un certain nombre de valeurs sémantiques (le syntagme réfère-t-il à un L&D ? si oui quelle est sa polarité, etc).

À ce niveau, ainsi qu'au niveau supérieur, l'analyse se veut à la fois syntaxique et sémantique. Suivant une approche compositionnelle, nous analysons la manière dont la valeur propre à chaque lexie contribue à la valeur sémantique du syntagme, selon sa catégorie grammaticale et la nature de la relation sémantique qu'elle entretient avec les autres lexies. Ainsi, en présence d'un énoncé comme "I don't like being indoors" (voir la figure 6.5), le système identifie trois syntagmes : un syntagme nominal, composé du pronom "I", un syntagme verbal, composé de la marque de négation "don't", et une proposition infinitive, "being indoors". À l'intérieur du syntagme verbal, il est en mesure de définir que la négation porte sur le verbe, et que dans ce cas la règle d'inversion de polarité doit être appliquée.

Notre définition du groupe syntaxique est ici à envisager de manière plus opérationnelle que théorique. Nous intégrons en effet autant des éléments proprement syntagmatiques que des éléments de nature propositionnelle. À ce niveau, l'analyse est ainsi en mesure de détecter tant des groupes minimaux comme les syntagmes adverbiaux que des groupes plus larges comme des propositions relatives ("The movie **that I saw last night** was terrible") ou non-finies, infinitive ("It was stupid to do that") ou participiale ("I love **playing tennis**"). Au total, huit catégories grammaticales sont envisagées : syntagme adverbial (*SAdv*), syntagme adjectival (*SAdj*), syntagme nominal (*SN*), syntagme prépositionnel (*SP*), syntagme verbal (*SV*), proposition relative (*REL*), proposition infinitive (*INF*) et proposition participiale (*PART*).

Lorsqu'une lexie de type L&D est détectée dans le syntagme analysé, le système considère le syntagme comme référant à un L&D. Pour ce type de syntagme, nous utilisons la représentation formelle *syntagmeCategorie(L&D)*, par exemple : *SN(L&D)*, *SAdj(L&D)*, *SV(L&D)*. La polarité est alors définie selon les règles décrites dans le chapitre 3. Chacune d'elles s'appliquent selon le type d'indice lexical permettant d'exprimer le L&D et selon la présence d'éventuels modificateurs ou de marques de négation à l'intérieur même du syntagme. Dans le cadre du syntagme, trois règles de polarité sont utilisées : la règle de neutralisation, la règle d'inversion et la règle d'agrégation. Pour cette première version du système, l'intensité n'a pas été traitée. L'ensemble des grammaires ainsi que les règles de polarité qu'elles intègrent sont détaillées dans l'annexe F.

Dans le cadre d'une application de négociation, le système traite en plus deux autres types de syntagmes : ceux référant à une volition (interaction uniquement de type négociation). Par exemple : *to have*, *doesn't want* ; et ceux qui réfèrent à un objet de la négociation, i.e à une cible potentielle (interaction uniquement de type négociation). Par exemple : *the art deco artwork*.

L'identification des syntagmes se fait grâce à des grammaires locales implémentées sous formes d'automates à états finis via la plateforme *Unitex*. Toutes les grammaires de l'analyse syntagmatiques sont appliquées de manière ascendante et selon un ordre précis sur le tour de parole. Ainsi, les grammaires permettant de reconnaître des groupes de niveau supérieur

peuvent exploiter des informations issues de l'application des grammaires de niveau inférieurs. Par exemple, la grammaire permettant de reconnaître un syntagme nominal, exploitent les informations issues de l'application de la grammaire dédiée aux groupes adjectivaux et adverbiaux. Chacune de ces grammaires formalise plusieurs réalisations syntaxiques pour une même catégorie de syntagme. Le tableau 6.2 ci-dessous présente les principales grammaires utilisées, le type de syntagme qu'elles détectent, le label qu'elles produisent et le nombre de structures qu'elles peuvent reconnaître.

Catégorie Grammaticale	Label	Exemples	Structures
Syntagme adverbial	<i>SAdv</i>	<i>very much</i>	1
	<i>SAdv(L&amp;D)</i>	<i>unfortunately</i>	1
Syntagme adjectival	<i>Sadj</i>	<i>very tall</i>	2
	<i>SAdj(L&amp;D)</i>	<i>scarely beautiful</i>	2
Syntagme nominal	<i>SN</i>	<i>The blue chair</i>	6
	<i>SN(L&amp;D)</i>	<i>A beautiful painting</i>	6
Syntagme prépositionnel	<i>SP</i>	<i>at home</i>	2
	<i>SP(L&amp;D)</i>	<i>with a beautiful color</i>	2
Syntagme verbal	<i>SV</i>	<i>is swimming</i>	17
	<i>SV(Attr)</i>	<i>isn't</i>	18
	<i>(SV(L&amp;D))</i>	<i>has been terrified</i>	16
Proposition infinitive	INF	<i>to watch this movie</i>	5
Proposition relative	REL	<i>the movie that i prefer</i>	4
Proposition participiale	PART	<i>Going to the park</i>	3

TABLEAU 6.2 – Ensemble des groupes syntaxiques traités

### 6.2.3 Analyse phrastique

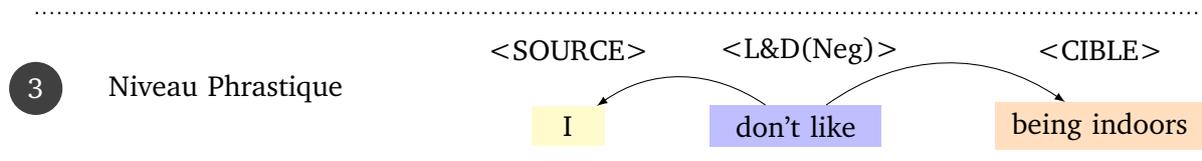


FIGURE 6.6 – Analyse phrastique du tour de parole de l'utilisateur

L'objectif de l'analyse phrastique est d'identifier des expressions de L&D à partir des blocs constitués au cours de l'analyse précédente. Lorsque l'analyse phrastique détecte une expression de L&D, elle identifie la source, la cible et détermine sa polarité sur la base de la relation logico-sémantique existant entre les différents syntagmes. À ce niveau, l'analyse est à la fois syntaxique et logico-sémantique : syntaxique, parce qu'elle cherche à définir les limites de la phrase et à identifier sa structure de surface, et logico-sémantique, parce qu'elle identifie les rôles sémantique des principaux syntagmes afin de pouvoir identifier la source et la cible du L&D. Pour plus de clarté, nous présentons ces traitements de manière successive, mais les grammaires les exécutent de manière simultanée.

**Aspects syntaxiques** L'analyse phrastique exploite des grammaires locales permettant de vérifier si les énoncés considérés correspondent dans leur structure syntaxique à des expressions de L&D. Les grammaires ont été définies sur la base des patrons décrits dans le chapitre 3. En tout 7 structures sont envisagées, 2 structures prédicatives et 5 structures attributives (voir Tableau 6.3). Les 2 structures prédicatives impliquent un verbe principal, qui joue le rôle de prédicat et réfère à un L&D (leur différence repose sur le type de structure actancielle qu'elles impliquent et que nous détaillons ci-dessous). Ainsi la phrase (i) est reconnue comme une expression de L&D dont la structure correspond au patron (ii). Toutes ces structures impliquent au moins un syntagme de L&D ( $SAdj(L&D)$ ,  $SN(L&D)$ ,  $SV(L&D)$ ). Pour chacune d'elles, une grammaire spécifique a été développée. Comme pour les syntagmes, chacune des grammaires du niveau phrastique est capable de reconnaître pour un même patron plusieurs réalisations syntaxiques (voir l'annexe F).

- (i) I don't like playing tennis
- (ii)  $SN + SV_{Type1}(L&S) + PART$

**Aspects logico-sémantiques** Lorsqu'une correspondance est trouvée entre l'énoncé analysé et l'une des structures représentées par les grammaires, l'identification de la source et de la cible peut alors être effectuée.

Dans les structures de type attributif, l'identification dépend de la nature du syntagme occupant la place de sujet : s'il s'agit d'un pronom personnel de la première personne, on le considère comme étant la source, autrement on le considère comme la cible. Dans le premier cas, le rôle de cible est attribuée au groupe rattaché à l'attribut, dans le second cas la source est assimilé au locuteur, en ce qui nous concerne l'utilisateur donc. Par exemple, dans l'énoncé (i) ci-dessous, le rôle de sujet est occupé par un pronom personnel de la première personne ("I"), nous considérons qu'il s'agit de la source. La cible est alors à chercher du côté de la proposition infinitive *to see the movie*. Dans l'exemple (ii) en revanche, le sujet est considéré comme la cible, la source est considérée comme non-exprimée syntaxiquement. L'attitude est donc attribuée au locuteur, i.e l'utilisateur.

- (i) "I was happy to see the movie"
- (ii) "He seems to be a bad guy"

Structures attributives	
$(SN1 INF PART) + SV(Attr) + SAdj(L\&D)$	“This painting is beautiful”
$(SN1 INF PART) + SV(Attr) + SN2(L\&D)$	“This painting is a masterpiece”
$(SN1 INF PART) + SV(Make) + SN2 + SV2 + SAdj(L\&D)$	“This movie makes me feel happy”
$(SN1 INF PART) + SV(Opinion) + SN2 + SV(Attr) + SN3$	“I think the painting is a masterpiece”
$(SN1 INF PART) + SV(Opinion) + SN2 + SV(Attr) + SAdj$	“I think the painting is beautiful”
Structures prédicatives	
$SN1 + SV_{Type1}(L\&S) + (SN1 INF PART)$	“I don’t like being indoors”
$(SN1 INF PART) + SV_{Type2}(L\&S) + SN1$	“Being indoors annoys me”

TABLEAU 6.3 – Structures syntaxiques des *likes* et *dislikes* communes aux applications *small-Talk* et *négo-ciation* (tour de parole de l'utilisateur)

Au niveau phrastique, des règles sémantiques de calcul de polarité doivent également être appliquées. Trois règles peuvent être appliquées selon les mécanismes décrits au chapitre 3 : inversion, agrégation et neutralisation. Leur application est décidée sur la base d'indices morpho-syntaxiques (la présence de modifieurs par exemple). Directement intégrées aux grammaires, elles sont appliquées consécutivement à l'identification d'une structure syntaxique pertinente.

## 6.3 Paire adjacente : détection des *likes* et *dislikes* strictement collaboratifs

Après une analyse du tour de parole de l'utilisateur, le système élargit le cadre d'analyse pour considérer la paire adjacente dans sa globalité. Cette phase de l'analyse se concentre exclusivement sur les expressions de *likes* et *dislikes* strictement collaboratives (L&DSC). Pour cela, le système analyse conjointement l'énoncé de l'agent et celui de l'utilisateur. En ce qui concerne l'agent, le système repère des expressions référant à un *like/dislike* pouvant être validées ou invalidées par l'utilisateur. Puisque nous ne travaillons pas avec un moteur de dialogue générant automatiquement les énoncés de l'agent, cette analyse est nécessaire pour vérifier la présence de ce type d'expression et pour récupérer la valeur des attributs polarité, source et cible. En ce qui concerne l'utilisateur, il doit déterminer si la réponse équivaut à un désaccord, invalidant le L&D de l'agent, ou à un accord, validant le L&D de l'agent et permettant à l'utilisateur d'en assumer également la responsabilité. Cette analyse est commune dans son architecture aux contextes *small-talk* et négociation.

### 6.3.1 Tour de parole de l'agent

Pour définir si un *like* ou *dislike* strictement collaboratif (L&DSC) est exprimé au niveau de la paire adjacente, il est nécessaire de savoir si l'agent lui-même exprime un L&D sous forme de question (exemple (i)) ou d'affirmation (exemple (ii)).

- (i) "Do you like outdoors activities?"
- (ii) "I hate Christmas"

En l'état actuel du processus de génération des énoncés de l'agent, il n'est pas possible de récupérer automatiquement cette information *via* le moteur de dialogue. Nous avons donc dû mettre en place une chaîne de traitement permettant de détecter ce type d'énoncé dans le tour de parole de l'agent. À la manière du traitement du tour de parole de l'utilisateur, cette chaîne de traitement repose sur une analyse ascendante considérant successivement trois cadres : un cadre lexical, un cadre syntagmatique et un cadre phrastique. L'analyse lexicale utilise les mêmes ressources que pour l'analyse du tour de parole de l'utilisateur. Elle a pour but de détecter des indices lexicaux de L&D et, dans le cadre d'application à une interaction de type négociation, des indices référant à un acte de volition et aux objets de la négociation. L'analyse syntagmatique est identique à celle menée sur le tour de parole de l'utilisateur : l'objectif est de détecter des syntagmes pouvant ensuite former des expressions de L&D. Au niveau phrastique, le système analyse le contexte d'emploi des syntagmes préalablement détectés pour vérifier s'ils correspondent dans leur structure à une expression de L&D. Deux formes sont envisagées : les formes affirmatives et les formes interrogatives. Pour les premières, les mêmes structures que pour l'analyse du tour de parole de l'utilisateur sont prises en compte. Pour les secondes, listées dans le tableau 6.4, il s'agit en fait de la forme interrogative des premières. Au terme de l'analyse, une valeur est attribuée aux variables  $Assertion(L\&D) = [True|False]$  et  $Question(L\&D) = [True|False]$ . La valeur des attributs *polarité* ou orientation de polarité pour les questions ouvertes (voir le chapitre 4), *source* et *cible* est déterminée sur la base des relations de dépendance définies dans le cadre de l'analyse phrastique.

Au terme de l'analyse, si une question ou une assertion véhiculant un L&D ( $Question(L\&D)$  ou  $Assertion(L\&D)$ ), le système lui attribue une structure de traits (représentée dans la seconde partie du tableau 6.4) définissant la valeur des attributs polarité/orientation polarité, source et cible (se référer au chapitre 4 pour une explication détaillée de ces attributs).  $T$  représente l'ensemble des entités qui peuvent être cibles d'un *like* ou d'un *dislike*. Dans une interaction de type *small-talk*, celles-ci sont relatives aux thèmes de la conversation,  $T$  ne peut donc être connu *a priori* et *in extenso*. En revanche, dans une application de type négociation,  $T$  correspond à l'ensemble des objets en jeu pendant la négociation.

Structures attributives	
$AUX + SN1 + SV(Make) + SN2(ProP2) + SV2 + SAdj(L\&D)$	
“Does this movie make you feel happy?”	
$AUX + SN1(ProP2) + SV(Opinion) + SN2 + (SV(Attr) as) + [SN3(L\&D) SAdj(L\&D)]$	
“Do you think the painting is beautiful?”	
Structures prédicatives	
$AUX + SN1(ProP2) + SV_{Type1}(L\&S) + (SN2 INF PART)$	
“Do you like being indoors?”	
Valeurs sémantiques associées au tour de parole de l’agent	
$Assertion(L\&D)$ Agent : “I like this painting”	$Polarity_A = [pos, neg]$ $Target_A \in T$ $Source_A = \{Agent, User\}$
$Question(L\&D)$ Agent : “Do you like this painting”	$Orientation_A = [pos, neg]$ $Target_A \in T$ $Source_A = \{Agent, User\}$

TABLEAU 6.4 – Structures syntaxiques et valeurs sémantiques des L&amp;D (tour de parole de l’agent)

### 6.3.2 Tour de parole de l’utilisateur

Si le système atteste de la présence d’une question ou d’une assertion portant sur un L&D dans le tour de parole de l’agent

$$Question(L\&D) == True \text{ OR } Assertion(L\&D) == True$$

alors la recherche d’une réponse de type accord ou désaccord est effectuée. Sur le plan sémantique, plusieurs configurations sont possibles (voir le tableau 6.5) :

- si l’assertion ou la question *ne contiennent pas* de marque de négation et que le système peut détecter un segment équivalent à un *yes*, alors la réponse est considérée comme un **accord** ;
- si l’assertion ou la question *ne contiennent pas* de marque de négation et que le système peut détecter un segment équivalent à un *no*, alors la réponse est considérée un **désaccord** ;
- si l’assertion ou la question *contiennent* une marque de négation et que le système peut détecter un segment équivalent à un *yes*, alors la réponse est considérée un **désaccord** ;
- si l’assertion ou la question *contiennent* une marque de négation et que le système peut détecter un segment équivalent à un *no*, alors la réponse est considérée un **accord** ;

### 6.3.3 Calcul des propriétés sémantiques

Selon les valeurs sémantiques définies lors de l’analyse du tour de parole de l’agent et celles de la réponse de l’utilisateur, le système peut déterminer si une L&DSC est exprimée au sein de la paire adjacente courante. Pour définir la valeur des attributs source, polarité et cible, les règles

Tour de parole agent	Tour de parole utilisateur
$Negation([Question Assertion]) == True$ “So, you don’t like Picasso?”	$Yes == True :$ “Yes, I do” Answer = Disagreement
	$No == True :$ “No” Answer = Agreement
$Negation([Question Assertion]) == False$ “Do you like Picasso”	$Yes == True :$ “Yes, I do” Answer = Agreement
	$No == True :$ “No, I don’t” Answer = Disagreement

TABLEAU 6.5 – Calcul de la valeur sémantique de la réponse de l'utilisateur

définies ci-dessous, ont été conçues comme des implémentations de la formalisation présentée dans le chapitre 4. Ainsi, ce sont les valeurs des variables  $Assertion(L\&D) = [True|False]$ ,  $Question(L\&D) = [True|False]$  et  $Answer = [Accord|Desaccord]$ , qui sont à la base du calcul de la valeur sémantique de ces attributs (tableau 6.6).

Tour de parole agent	Tour de parole utilisateur
$Assertion(L\&D) == True$ AND $Answer == Agreement$	$Polarity(L\&D) == Polarity_A$ $Target(L\&D) == Target_A$ $Source(L\&D) == User$
$Assertion(L\&D) == True$ AND $Answer == Disagreement$	$Polarity(L\&D) == Reverse(Polarity_A)$ $Target(L\&D) == Target_1$ $Source(L\&D) == User$
$Question(L\&D) == True$ AND $Answer == Agreement$	$Polarity(L\&D) == Orientation_A$ $Target(L\&D) == Target_A$ $Source(L\&D) == User$
$Question(L\&D) == True$ AND $Answer == Disagreement$	$Polarity(L\&D) == Reverse(Orientation_A)$ $Target(L\&D) == Target_A$ $Source(L\&D) == User$

TABLEAU 6.6 – Calcul de la valeur sémantique des attributs polarité, cible et source de la L&DSC détectée

## 6.4 Séquence thématique : optimiser le traitement des cibles

Les interactions humain-agent sont structurées par un scénario qui guide la progression de la conversation. Ce scénario offre la possibilité de connaître les thèmes qui pourront être abordés au cours de la conversation. L'objectif des analyses menées à ce niveau est d'ancrer la détection de *likes* et de *dislikes* dans la structure thématique, afin notamment d'améliorer le traitement des cibles. Pour cela, cette phase de l'analyse a été conçue pour exploiter des informations fournies par le moteur de dialogue concernant la nature des thèmes introduits par l'agent. Cet ancrage thématique s'organise en trois étapes. Tout d'abord, nous avons procédé à une mise en relation entre les thèmes impliqués par le scénario et les catégories de domaine contenus dans *WordNet-Domain* (niveau A). Ensuite, le système procède à une détection des différentes lexicalisations du thème (i.e. les mots référant au thème ou à l'un de ses sous-aspects), en exploitant les relations préalablement posées entre *WordNet-Domain* et les thèmes du scénario (niveau B). Enfin, le système effectue un rapprochement sémantique entre les cibles et les thèmes de la conversation (niveau C). L'ensemble des traitements présentés dans cette section est effectué pour toute séquence thématique clôturée. Ils s'appliquent sur l'ensemble des *likes* et *dislikes* détectés au cours des analyses du tour de parole (section 6.2) et de la paire adjacente (section 6.3) présentées dans les sections précédentes. Ce niveau de l'analyse n'a pas été développé pour le contexte applicatif de la négociation.

### 6.4.1 Niveau A : relations entre les thèmes et les domaines de *WordNet-Domains*

Nous avons identifié 7 thèmes génériques pouvant être ouverts par l'agent : *Free-Time Activities*, *Free-Time Projects*, *Professional Activity*, *Professional Projects*, *Generic Projects*, *Happiness*, *Anger*. Cette cartographie des thèmes possibles n'étant pas livrée avec le scénario *Semaine*, nous l'avons déduite de l'observation du corpus de travail et des questions qui y sont posées à l'utilisateur. Nous considérons ces thèmes comme génériques car ils fournissent à l'utilisateur un cadre au sein duquel il pourra définir un thème spécifique. Dans l'exemple (i) ci-dessous, la question posée par l'agent ouvre le cadre thématique *Free-Time Projects* et la réponse de l'utilisateur instancie dans ce cadre le thème *Islands*. Dans l'exemple, l'agent ouvre le cadre thématique *Happiness* au sein duquel l'utilisateur instancie le thème *Christmas*.

(i) Agent : "What have you plan for your vacation?"

(i) User : "I want to go to Santorini. It's a such beautiful island. I love islands"

(ii) Agent : "Tell me about the happy things that have happened to you"

(ii) User : "I had a nice Christmas"

Si l'on souhaite ancrer les résultats de la détection des L&D (simples et strictement collaboratifs) dans la structure thématique, il est nécessaire de pouvoir identifier le thème spécifique instancié par l'utilisateur. Or, du point de vue de la détection, le problème est que celui-ci n'est pas connu *a priori* et ne peut être identifié que par une analyse du contenu verbal de l'utilisateur. Pour rendre possible cette identification, nous procédons à une analyse lexicale aboutissant à l'identification de mots-thèmes (cette analyse est expliquée ci-dessus, niveau B). La liste de mots-thèmes ne permet pas de nommer le thème, mais permet d'avoir un aperçu de sa lexicalisation et de son contour sémantique.

Afin de rendre possible cette extraction des mots-thèmes, nous avons rattaché chaque cadre thématique à un ou plusieurs domaine de *WordNet-Domains*. Comme expliqué précédemment (Section 6.2.1), *WordNet-Domain* associe ressources lexicale et conceptuelle : cette ressource définit une hiérarchie conceptuelle de domaines auxquels sont ensuite rattachés des *synsets* issus de *WordNet*. Pour cette première exploration, les relations sont été établies manuellement. Neuf domaines ont été considérés comme pertinents : *Free-time*, *Meteorology*, *Gastronomy*, *Play*,

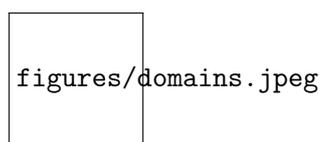


FIGURE 6.7 – Liens entre les thèmes identifiés dans le corpus *Semaine* et les domaines de *WordNet-Domain*

*Geography, University, Pedagogy*. L'ensemble des relations établies est illustré à la figure 6.7 : les ellipses représentent les thèmes *Semaine*, les termes en italique, les domaines.

### 6.4.2 Niveau B : détection de mots-thèmes

Les mots-thèmes sont les mots qui, dans le contenu verbal de l'utilisateur, sont associés sur le plan sémantique au *thème* courant de la conversation et participent à l'élaboration de son isotopie sémantique. Ces mots sont détectés dans chaque tour de parole de l'utilisateur intégrant la séquence thématique courante. Le processus de détection prend en compte l'ensemble des noms employés par l'utilisateur, le thème générique ouvert par l'agent, les domaines associés au thème et les synsets associés à chacun de ces domaines. Formellement, ces entrées sont représentées comme suit :

- $Topic_{Agent} \in T$  où  $T$  est l'ensemble des thèmes génériques définis sur la base du scénario,  $T = \{FreeTimeActivities, FreeTimeProjects, ProfessionnalActivity, ProfessionnalProject, Project, Happiness, Anger\}$
- $RelevantDomains = \{domain_1, domain_2, \dots, domain_n\}$  : un domaine appartient à l'ensemble des *RelevantDomains*, si il a été associé à l'un des  $TopicFrame_{Agent}$  ;
- $Nouns_{User} = \{noun_1, noun_2, \dots, noun_n\}$  : la liste des noms employés par l'utilisateur.
- $SynsetInDomains = \{synset_1, synset_2, \dots, synset_n\}$  : la liste des synsets rattachés à chaque domaine.

Pour chaque  $nom_i \in Nouns_{User}$ , le système crée une liste  $SynsetList(noun_i)$ , i.e. une liste de tous les *synsets* comprenant le  $noun_i$ . Pour chaque *synset*, si le *synset* est rattaché à un domaine, on considère alors que le nom est un mot topic : si  $synset \in SynsetInDomain$ , alors  $noun_i = TopicWord$  (i.e un mot-thème). Chaque *TopicWord* détecté est ajouté à la liste  $TopicWordsList$  associée à la séquence de thématique traitée. Le système peut prendre en charge pour chaque séquence plusieurs domaines, néanmoins ils doivent être inclus dans *RelevantDomains*.

### 6.4.3 Niveau C : vers un modèle utilisateur

Lorsque la détection de mots-thèmes est effectuée, pour toute séquence thématique, nous disposons d'une liste des *likes* et *dislikes* ( $L\&DList$ ), et une liste de mots-thèmes ( $TopicWdList$ ). La création du modèle utilisateur implique une organisation et une structuration sémantique des résultats. La création du modèle s'organise en trois temps. Dans un premier temps, le système analyse les cibles de tous les L&D trouvés dans la séquence thématique afin de fusionner les expressions ayant des cibles similaires et de la sorte supprimer les doublons. Dans un second temps, le système cherche à établir des relations sémantiques entre les cibles et le thème courant. Enfin, dans un troisième temps, le système applique une heuristique simple pour résoudre *a minima* les problèmes de coréférence.

#### Niveau C1 : fusion des cibles

Les *likes* et *dislikes* détectés dans l'ensemble de la séquence ont tous pour cible syntaxique un syntagme dont la catégorie varie selon les cas (relative, infinitive, syntagme nominal). Afin de fusionner les cibles référant au même concept, le système extrait le nom-tête du syntagme

défini comme cible syntaxique. Il obtient ainsi une liste de noms-cibles ( $Noun(Target)$ ). Dans un second temps, il procède à une comparaison de tous les noms-cibles ainsi trouvés et tente de déterminer s'ils sont similaires ou synonymes. La synonymie est établie par une analyse de leurs *synsets* respectifs. Si les deux mots cibles partagent le même *synset*, alors la relation de synonymie peut être établie. Au final, toute relation de synonymie ou de similarité trouvée entre les noms-cibles de deux expressions de L&D ayant la même polarité, indique que les deux expressions sont deux occurrences référant au même *like* ou *dislike*.

Formellement, cette analyse peut être représentée comme suit :

Similarité Lexicale :

```
if  $Pol_{LikeOrDislike_i} == Pol_{LikeOrDislike_j}$ 
and  $Noun_{Target_i} == Noun_{Target_j}$ ,
then  $Similarity(Noun_{Target_i}, Noun_{Target_j}) = True$ 
```

Synonymie :

```
if  $Pol_{LikeOrDislike_i} == Pol_{LikeOrDislike_j}$ 
and  $ShareSynset(Noun_{Target_i}, Noun_{Target_j})$ ,
then  $Synonymy(Noun_{Target_i}, Noun_{Target_j}) = True$ 
```

```
Fusion : if  $Synonymy(Noun_{Target_i}, Noun_{Target_j}) == True$ 
or  $Similarity(Noun_{Target_i}, Noun_{Target_j}) == True$ 
then  $Target_i = Target_j$ 
```

## Niveau C2 : correspondance entre les cibles et le thème courant

Nous posons ici l'hypothèse qu'une cible est toujours sémantiquement liée au thème courant. Par défaut, nous établissons une relation sémantique entre la cible et le thème. Au-delà de cette relation par défaut, le système cherche à préciser la nature de cette relation. Trois relations différentes sont prises en compte : une relation de similarité, une relation de synonymie et une relation de proximité d'aspect. Pour définir la nature de cette relation, nous comparons chaque nom-cible identifié au niveau C1 ( $Noun(Target)$ ) à chaque mots-thème *TopicWord* identifié au niveau B. Si le nom-cible et le mot-thème sont identiques, il est possible de conclure à une relation de similarité. Au cours de l'analyse des *synsets* du mot-thème et du nom-cible, si nous pouvons vérifier que ceux-ci partagent le même *synset* alors une relation de synonymie peut être établie. En cas d'absence de similarité ou de synonymie, nous posons l'hypothèse que le nom-cible réfère à l'une des facettes du thème courant (similarité d'aspect). Les relations de méronymie/holonymie et hyperonymie/hyponymie n'ont pas été traitées dans le cadre de cette première version du système.

Formellement, l'analyse peut être représentée comme suit :

Similarité

```
if  $Noun_{Target} == TopicWord$ ,
then  $Similarity(Noun_{Target}, TopicWord) == True$ 
```

Synonymie

```
if  $ShareSynset(Noun_{Target}, TopicWord)$ ,
then  $Synonymy(Noun_{Target}, TopicWord) == True$ 
```

Proximité d'aspect

```
if  $Similarity(Noun_{Target}, TopicWord) == False$ 
and  $Synonymy(Noun_{Target}, TopicWord) == False$ ,
then  $AspectProximity(Noun_{Target}, InstanciatedTopic)$ 
```

### Niveau C3 : gestion *a minima* de la coréférence

Lorsque le syntagme identifié comme cible syntaxique de l'expression est en fait un pronom indéfini ou démonstratif, la cible est considérée comme inconnue.

```
If Target == [it|this|that|which]  
then Target = unknownTarget
```

La solution proposée pour intégrer malgré tout ce type d'expression de L&D au modèle utilisateur, est de considérer les cibles inconnues comme des facettes du thème courant. Ainsi, puisqu'aucune information sémantique ne peut être obtenue à propos de ce type de cibles, nous rattachons celles-ci directement au thème.

### 6.4.4 Exemple sur une séquence du corpus Semaine

**Niveau A : détection des *likes* et *dislikes* (figure 6.8)** Après l’analyse de l’ensemble des tours de parole de l’utilisateur et de l’ensemble des paires adjacentes dans la séquence thématique, le système détient la liste de tous les L&D exprimés au cours de la séquence. La figure ?? représente une séquence thématique issue de la session 21 du corpus Semaine. Les tours de parole notés  $O_n$  sont ceux de l’opérateur, et ceux notés  $U_n$  sont ceux de l’utilisateur. Nous rappelons ici que les énoncés de l’opérateur ne sont pas générés automatiquement mais contraints par un script que doit suivre l’opérateur jouant le rôle de l’agent. Ceux de l’utilisateur sont en revanche spontanés. Au total 7 expressions de L&D ont été détectées par le système, elles sont présentées dans la colonne de droite de la figure ?? . La polarité a été correctement détectée et le syntagme correspondant à la cible syntaxique a été correctement identifié.

**Niveau B : détection des mots-thèmes (figure 6.8)** : Comme le montre la première phrase de la séquence, “So, tell me about happy things that have happened to you”, le thème générique défini par l’agent correspond dans notre cartographie de thèmes de Semaine, au thème *Happiness*. Après l’analyse des synsets rattachés aux *domaines* correspondants, le système identifie deux mots-thèmes : *Christmas, present*.

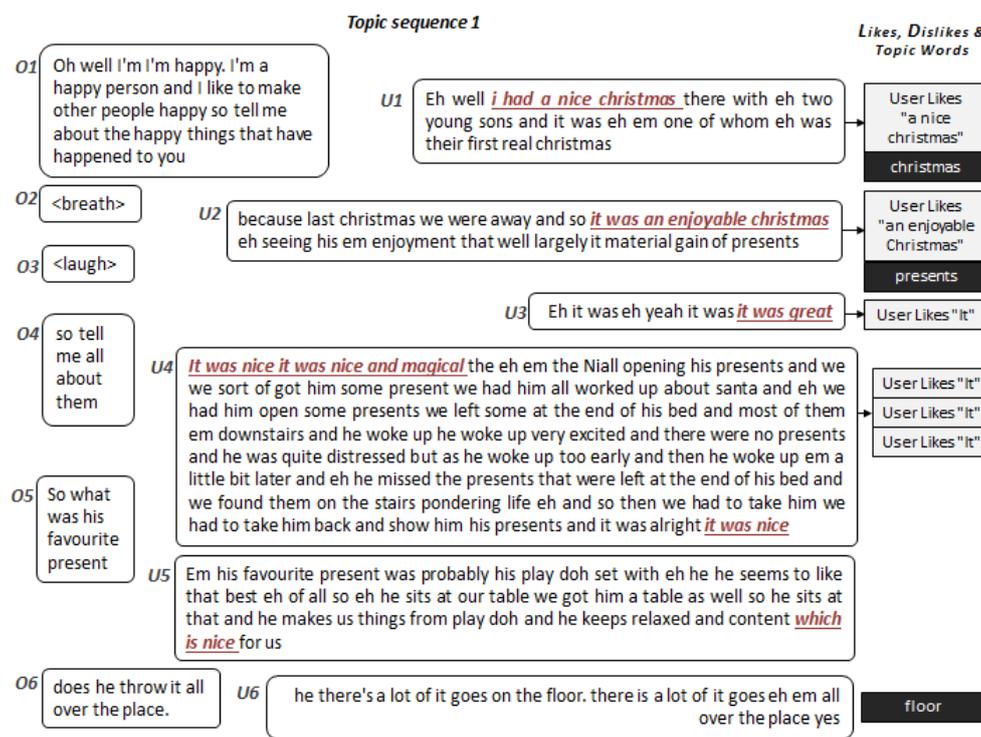


FIGURE 6.8 – Sorties des niveaux A et B

**Création du modèle utilisateur** : À ce niveau, le système dispose de deux listes : une liste de mots thèmes et une liste des *likes* et *dislikes*.

- Fusion des cibles (Figure 6.9). 6 expressions de *likes* sont détectées dans la séquence. Les deux premières ont pour cible “a nice Christmas” et “an enjoyable Christmas”. Ici, l’analyse des cible permet de les considérer comme deux lexicalisations de la même cible concept. Les deux expressions ayant ainsi la même cible et la même polarité, elles sont fusionnées. Les 4 autres expressions ont pour cible syntaxique, le pronom indéfini *it*.

Celui-ci réfère à un mot que le système n'a pas identifié. Les 4 expressions sont regroupées au sein d'un même ensemble de *likes* dont la cible est notée comme inconnue.

- Correspondance entre cibles et thèmes : au terme de l'opération de fusion des cibles, une seule cible est candidate pour l'analyse du niveau analyse de correspondance sémantique entre le thème et les cibles. Suite à l'analyse de relations sémantiques possibles, une relation de *similarité* est détectée.
- Gestion de la coréférence : à ce niveau, le système prend en charge les expressions dont les cibles sont inconnues. Pour résoudre *a minima* ce problème de coréférence, le système établit une relation entre la cible et le thème courant, en considérant la dernière comme une facette du premier.

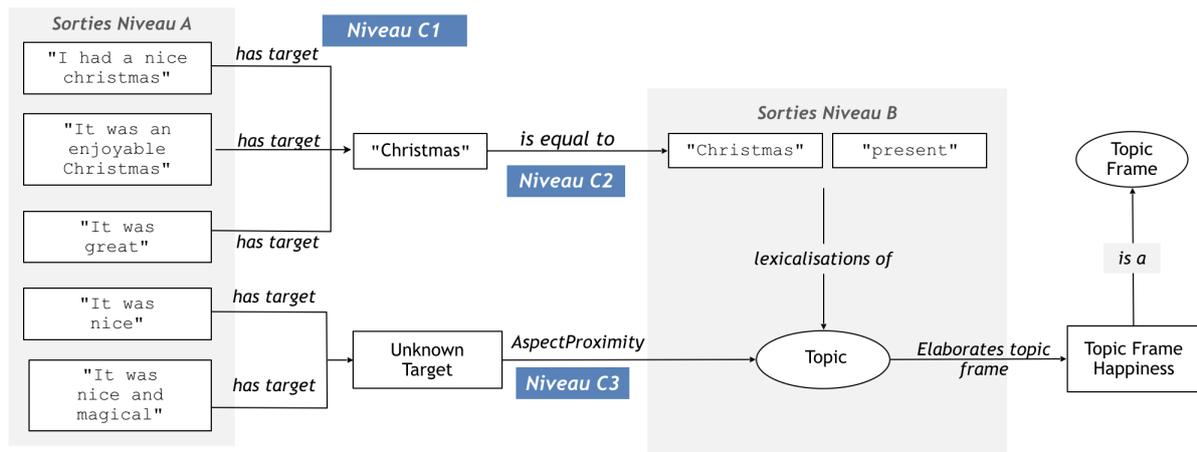


FIGURE 6.9 – Analyses du niveau C

## 6.5 Conclusion du chapitre 6

Dans ce chapitre, nous avons présenté un modèle de détection des *likes* et *dislikes* de l'utilisateur au cours de conversations avec un agent conversationnel animé. Ce modèle se concentre sur deux types d'interaction : des conversations de *small-talk* et des sessions de négociation. La constitution d'une catégorie de *likes* et *dislikes* est la conséquence de la prise en compte d'un impératif applicatif : dans le cadre conversationnel considéré, ce type d'expression constitue un atout pour l'agent et pour l'élaboration de stratégies dialogiques et interactionnelles. Cette catégorie doit être comprise comme une catégorie opérationnelle permettant de caractériser du point de vue de l'agent les énoncés pertinents pour la détection. Elle ne s'oppose pas à la catégorisation des *Attitudes* de Martin and White (2005), que nous envisageons comme une catégorisation théorique utile à la description linguistique du phénomène. Le modèle des attitudes permet de définir les caractéristiques distinctives d'un ensemble d'expressions de nature axiologique et affective, sur la base desquelles il nous a été possible de sélectionner celles qui nous paraissent les plus pertinentes et les plus à même de donner des indices sur les goûts de l'utilisateur.

Au cours de la conversation, le système considère trois cadres d'analyse : le tour de parole, la paire adjacente et la séquence thématique. À chacun de ces niveaux, le système se concentre sur un sous-phénomène spécifique.

- Au niveau du tour de parole, l'objectif est de détecter des expressions simples de L&D, i.e des expressions dont l'interprétation repose sur des éléments verbaux contenus au sein même du tour de parole. L'analyse se déroule selon un processus ascendant considérant successivement le lexique, les syntagmes et la structure syntaxique et logico-sémantique de la phrase. Ce processus d'analyse ascendant exploite des grammaires formelles combinant des règles sémantiques à des patrons syntaxiques.
- Au niveau de la paire adjacente, le système se concentre sur le caractère participatif des expressions de L&D et plus spécifiquement un sous-aspect de cette caractéristique, les expressions de L&D strictement collaboratives. À la différence des expressions simples, l'interprétation des expressions strictement collaboratives repose sur la prise en compte d'éléments sémantiques et pragmatiques repartis sur les deux tours de parole qui composent la paire adjacente. L'analyse à ce niveau combine traitement du tour de parole de l'agent et traitement du tour de parole de l'utilisateur. L'énoncé de l'agent est traité selon un processus d'analyse ascendant similaire à celui déployé au niveau précédent. Lorsqu'une expression de L&D est détectée chez l'agent, le système cherche la présence d'un accord ou d'un désaccord dans la réponse de l'utilisateur. Selon les éléments ainsi détectés, la valeur des attributs *cible*, *source* et *polarité* peut être définie.
- Au niveau de la séquence thématique, afin de produire un modèle utilisateur, le système a pour objectif d'ancrer les résultats produits par les deux précédents niveaux à la structure thématique de la conversation. Pour cela, il prend en compte les informations produites par le moteur de dialogue concernant le thème introduit par l'agent. Sur la base de cette information, il exploite la ressource *WordNet-Domains* pour détecter dans les énoncés de l'utilisateur l'ensemble des lexicalisations (mots-thèmes) du thème courant. Une fois, les mots-thèmes rassemblés, il les compare aux cibles des L&D détectés pour tenter d'établir entre eux des relations de similarité ou de synonymie. Cette analyse ainsi que l'élimination de doublons, i.e de L&D ayant la même polarité et la même cible, mais également d'une gestion *a minima* de la coréférence permettent de créer un modèle utilisateur structurant les résultats du système.

Les évaluations réalisées pour vérifier les performances de ces différents niveaux d'analyse sont présentées dans le chapitre suivant (chapitre7).

## Chapitre 7

# Annotation de corpus et protocoles d'évaluation

### Sommaire

---

7.1	Évaluation dans le cadre d'un scénario de <i>small-talk</i> (corpus Semaine) . . . . .	96
7.1.1	Évaluation <i>small-talk</i> 1 : tour de parole et paire adjacente . . . . .	96
7.1.2	Évaluation qualitative <i>small-talk topic</i> : séquence thématique . . . . .	100
7.2	Évaluation dans le cadre d'un scénario de négociation . . . . .	102
7.2.1	Protocole et plateforme d'annotation . . . . .	102
7.2.2	Accord inter-annotateur . . . . .	102
7.2.3	Évaluation et discussion . . . . .	103
7.3	Conclusion du chapitre 7 . . . . .	105

---

DANS ce chapitre, nous abordons la question de l'évaluation des différents aspects du système présenté dans le chapitre précédent, ainsi que celle de la constitution de jeux d'annotation nécessaires à cette évaluation. Les trois évaluations présentées dans ce chapitre s'intéressent aux performances des différents sous-niveaux d'analyse du système. L'évaluation *Small-Talk 1* et l'évaluation *Négociation* mesurent les performances du système au niveau du tour de parole et de la paire adjacente, sur le corpus *Semaine* pour la première et sur le corpus *Négociation* pour la seconde. L'évaluation *Small-Talk Topic* évalue les résultats produits par l'analyse des *likes* et *dislikes* dans le cadre de la séquence thématique. Les deux premières évaluations ont nécessité la mise en place de campagnes d'annotation via des plateformes de *crowdsourcing*. La troisième constitue davantage une évaluation qualitative se fondant sur une observation systématisée des résultats.

Les protocoles d'évaluation définis en traitement automatique du langage impliquent l'utilisation de corpus de références disposant du même jeu d'annotation que celui produit par le système. Trois modes de constitution sont généralement suivis pour créer ce type de ressources. Tout d'abord, il est possible de récupérer des textes issus du web incluant des annotations fournies par les internautes. Par exemple, le site *Rottentomatoes.com* a été utilisé pour créer des corpus de critiques de films destinés à l'entraînement et à l'évaluation de systèmes de classification ((Pang et al., 2008), (Pang and Lee, 2005)). L'avantage de telles ressources est que leur constitution ne nécessite pas le recrutement d'annotateurs. Néanmoins, elles peuvent demander un grand effort de nettoyage et de normalisation des données. De plus, les données et les annotations ainsi récupérées ne correspondent qu'à un certain type d'applications. Le second mode de constitution fait appel à des annotateurs experts, généralement des linguistes de formation. Pour ce type d'annotation, la création d'un schéma détaillé est nécessaire. Celui-ci permet de formaliser le phénomène à annoter et de donner des instructions précises aux annotateurs. Bien que coûteux, ce mode de constitution est choisi lorsque sont attendues des annotations à grain fin, impliquant des observations de grande précision des phénomènes linguistiques ciblés. Enfin, le troisième mode de constitution repose sur l'exploitation de plateformes de *crowdsourcing* pour le recrutement d'annotateurs. La plateforme *Amazon Mechanical Turk* a ainsi été utilisée pour la constitution de jeux d'annotation et l'évaluation de différents types d'application : évaluation de traduction automatique, validation de dictionnaire mais aussi reconnaissance automatique d'affect dans le langage. Pour notre travail, l'objectif n'était pas d'obtenir des annotations pour l'apprentissage d'un système de détection mais de vérifier les interprétations faites par les différentes chaînes de traitement. En ce sens, il nous a semblé moins nécessaire d'obtenir des annotations d'experts des phénomènes ciblés que de connaître la manière dont des locuteurs ordinaires les reconnaissent et les interprètent. Nous avons donc choisi de faire appel à des annotateurs non-experts. Pour la campagne *Small-Talk*, ces derniers ont été recrutés via *Amazon Mechanical Turk* (AMT), pour la campagne *Négociation*, via *CrowdFlower* (CF)<sup>1</sup>.

Une fois recrutés, les annotateurs étaient ensuite redirigés vers des plateformes dédiées aux tâches d'annotation. Deux plateformes ont été développées, une pour chaque évaluation. Ces plateformes, pensées pour une exploitation en ligne, supportent des schémas d'annotation qui partagent un certain nombre de caractéristiques :

- la mise en valeur du contenu verbal : les plateformes sont dédiées à l'annotation exclusive du contenu verbal. L'objectif étant d'obtenir des annotations considérant uniquement le contenu verbal sans l'influence du contexte multi-modal, les plateformes ne donnent accès qu'à la transcription des énoncés à annoter sans ajout d'information sur les indices non verbaux via l'audio ou la vidéo.
- la prise en compte de caractéristiques conversationnelles : afin de mettre en valeur la spécificité énonciative du contenu annoté, les deux plateformes mettent l'accent sur la paire adjacente. De cette manière, le contenu verbal de l'utilisateur peut être interprété par l'annotateur à la lumière de celui de l'agent. Afin de ne pas complexifier le processus

1. Le changement de plateforme, de AMT vers CF, n'a été dicté que par des impératifs administratifs

---

d’annotation, la distinction entre tour de parole et paire adjacente n’a pas été présentée aux annotateurs. Pour chaque paire adjacente considérée, l’annotateur peut identifier soit des L&D simples exprimés dans le tour de parole, soit des L&D strictement collaboratifs exprimés au niveau de la paire. Ainsi, il nous a été possible de récupérer des annotations correspondant aux annotations produites par notre système au niveau du tour de parole et de la paire adjacente.

- la simplification de la tâche d’annotation : contrairement à ce qui peut se faire dans de nombreuses plateformes d’annotation, généralement pensées pour des annotateurs experts, nous avons fait le choix de faire passer le processus d’annotation par des questionnaires. Ainsi, l’annotateur n’a pas à sélectionner des morceaux de texte, mais doit répondre à un certain nombre de questions.

Le présent chapitre s’organise en deux sections. La première concerne les deux évaluations menées sur le corpus *Semaine* permettant d’évaluer les performances du système dans le cadre d’une application de type *Small-Talk* : les évaluations *Small-Talk 1* et *Small-Talk Topic*. Pour *Small-Talk 1*, nous présentons la plateforme d’annotation, le protocole de recrutement sur AMT, les résultats du calcul de l’accord inter-annotateurs et ceux de l’évaluation du système, et une discussion de ces résultats. La seconde section est consacrée à la description de l’évaluation *Négociation* et aborde les mêmes sous-aspects : plateforme d’annotation, mesure d’accord inter-annotateurs, résultats de l’évaluation et discussion.

Publications associées à ce chapitre :

- Langlet, C. and Clavel, C. (2015b). Improving social relationships in face-to-face human-agent interactions: when the agent wants to know user’s likes and dislikes. In *Proceedings of the 53rd Annual Meeting of the Association for Computational Linguistics and the 7th International Joint Conference on Natural Language Processing (Volume 1: Long Papers)*, volume 1, pages 1064–1073
- Langlet, C. and Clavel, C. (2015a). Adapting sentiment analysis to face-to-face human-agent interactions: from the detection to the evaluation issues. In *Affective Computing and Intelligent Interaction (ACII), 2015 International Conference on*, pages 14–20. IEEE
- Langlet, C. and Clavel, C. (2016). Grounding the detection of the user’s likes and dislikes on the topic structure of human-agent interactions. *Knowledge-Based Systems*, 106:116–124
- Langlet, C., Duplessis, G. D., and Clavel, C. (2017). A web-based platform for annotating sentiment-related phenomena in human-agent conversations. In *International Conference on Intelligent Virtual Agents*, pages 239–242. Springer
- Langlet, C. and Clavel, C. (2018). Detecting user’s likes and dislikes for a virtual negotiating agent. In *20th ACM International Conference on Multimodal Interaction (ICMI 2018)*, Boulder, Colorado, (Submit)

## 7.1 Évaluation dans le cadre d'un scénario de *small-talk* (corpus *Semaine*)

Cette section présente l'évaluation des performances du système dans le cadre d'une application de *small-talk*. Deux sous-aspects sont pris en compte : la détection d'expressions de *likes* et *dislikes* au niveau du tour de parole et de la paire adjacente, l'analyse des relations sémantiques entretenues entre la cible et le thème, au niveau de la séquence thématique. Ces deux évaluations ont été réalisées sur la partie de corpus *Semaine* non prise en compte pour le développement du système.

### 7.1.1 Évaluation *small-talk* 1 : tour de parole et paire adjacente

#### 7.1.1.1 Protocole et plateforme d'annotation

Le corpus d'évaluation consiste en un ensemble de paires adjacentes extraites du corpus *Semaine* (McKeown et al., 2011). Le corpus *Semaine* comprend 65 sessions transcrites manuellement dans lesquelles un utilisateur humain interagit avec un opérateur humain jouant le rôle de l'agent virtuel. Ces interactions sont basées sur un scénario impliquant quatre personnages agents : **Poppy**, joyeuse et extravertie, **Prudence**, sensée et pondérée, **Spike**, en colère et conflictuel et **Obadiah**, dépressif et morose. Les phrases de l'agent sont contraintes par un script (cependant, certaines déviations au script se produisent dans la base de données) visant à mettre l'utilisateur dans le même état que celui du personnage joué. Parmi les 35 sessions non utilisées pour le développement du système, nous avons sélectionné de manière aléatoire 600 paires ne dépassant pas un total de 30 mots. Cette limitation du nombre de mots nous a paru nécessaire pour simplifier une tâche d'annotation déjà complexe. En effet, les transcriptions de tours de parole excédant cette longueur sont souvent difficile à lire et à interpréter, du fait du nombre accru de disfluences. Chacune des paires adjacentes constituées contiennent un tour de parole de l'agent et un tour de parole de l'utilisateur. Les 600 paires ont été ensuite réparties en 60 sous-ensembles constitués de 10 paires. Les annotateurs ont été sélectionnés, sur AMT, en fonction du taux d'approbation (supérieur ou égal à 99%) – qui leur est attribué selon la qualité de leur travail d'annotation – et le nombre de tâches approuvées (supérieur ou égal à 10000). Au total, 240 annotateurs ont été recrutés. Chaque sous-ensemble du corpus a ensuite été attribué aléatoirement à 4 annotateurs.

Après leur recrutement sur AMT, les annotateurs ont été redirigés vers la plateforme développée par nos soins. Principe central de la conception de la plateforme, la mise en valeur du contenu a pour objectif de ne fournir aux annotateurs que la transcription textuelle sans les indices multimodaux et sans l'accès au contenu antérieur ou postérieur. De la sorte, l'entrée fournie aux annotateurs est identique à celle fournie au système et les annotations ainsi obtenues correspondent à une interprétation du contenu ne prenant en compte que le verbal. A chaque tour d'annotation, la plateforme propose à l'annotateur une nouvelle paire adjacente. L'annotation s'effectue par l'intermédiaire de questionnaires. Selon la nature des réponses de l'annotateur, le nombre de questions varie. Ces différentes questions permettent de définir le nombre de *likes* et *dislikes* détectés par l'annotateur dans la paire adjacente présentée et de définir pour chacun d'eux la valeur et l'intensité de leur polarité, l'identité de leur cible (voir la figure 7.1). Afin de faciliter l'annotation et de rendre l'interprétation de chaque question aussi spontanée que possible, celles-ci ont été conçues sans termes techniques. Chaque question se concentre sur l'une des sorties du système de détection :

1. **Présence d'expressions de L&D** : la première question interroge la présence d'une expression de *like* ou de *dislike*. Aucune discrimination sur la polarité n'est faite à ce stade. Les questions suivantes ne s'appliquent que si une réponse positive à la première question a été donnée.

2. **Nombre d'expressions de L&D** : la deuxième question interroge le nombre d'expressions présentes dans une même paire adjacente. Nous avons limité la réponse à 4, ce qui correspond au nombre maximal d'expressions de *likes* ou *dislikes* trouvées dans le corpus de travail. Les questions 3a, 3b, 3c sont posées pour chaque expression trouvée dans la paire adjacente.
3. **Type de cible** : la question 3a interroge le type de la cible. Quatre différents types de cible sont proposés.
4. **Mots référant à la cible** : le formulaire 3b propose à l'utilisateur de sélectionner les mots qui, dans le tour de parole de l'utilisateur et de l'opérateur, réfèrent à la cible de l'expression traitée.
5. **Polarité** : la question 3c permet de définir la polarité (positive → *like*, négative → *dislike*) de chaque expression trouvée dans la paire adjacente.

Les données récupérées au terme de cette campagne d'annotation sont décrites dans l'annexe C.

### 7.1.1.2 Mesures d'accord inter-annotateurs

L'ensemble des mesures présentées dans cette section a été appliqué pour chaque sous-groupe d'annotateurs (c'est-à-dire chaque groupe de 4 annotateurs travaillant sur le même ensemble de 10 paires adjacentes).

	Kappa de Fleiss	Alpha de Cronbach
Maximal	0.79	0.90
Médiane	0.32	0.72
Moyenne	0.25	0.59

TABLEAU 7.1 – Scores du kappa de Fleiss et de l'alpha de Cronbach obtenus sur les 60 sous-ensembles (évaluation *small-talk* 1)

Pour la réponse à la première question (*Combien d'expressions de Likes ou Dislikes sont présentes dans cette paire adjacente?*), nous avons mesuré, dans un premier temps, le taux d'accord concernant la présence d'au moins un *like* ou *dislike*. Pour cela, nous avons utilisé le kappa de Fleiss (Fleiss, 1971) (Tableau 7.1). Dans un second temps, concernant le nombre de d'expressions identifiées, la mesure d'une relative similarité plutôt que celle d'un accord exact entre les chiffres donnés, nous a semblé plus pertinente. Pour cette raison, nous avons calculé la cohérence interne des réponses en utilisant le coefficient alpha de Cronbach (Cronbach, 1951). Cette mesure a été conçue pour évaluer l'homogénéité de scores donnés sur une échelle et ainsi le degré de cohérence avec lequel les annotateurs font des estimations sur le même phénomène. Les scores obtenus sont encourageants. Concernant l'accord sur la présence d'au moins une expression de *like* ou *dislike*, même si le score médian est compris entre 0.30 et 0.40, la valeur maximale obtenue est de 0,79. De plus 40% des sous-ensembles obtiennent un kappa compris entre 0.40 et 0.60. Les résultats obtenus pour le calcul de la cohérence montrent que 51% des sous-ensembles annotés ont un coefficient supérieur ou égal à 0.7, ce qui est considéré comme un niveau acceptable d'accord.

Pour la polarité et le type de cible, nous avons sélectionné les paires dans lesquelles au moins 2 annotateurs sont d'accord sur la présence d'une expression de *like* ou *dislike*, et nous avons ensuite considéré uniquement les réponses de ces annotateurs. Après cette sélection, nous avons

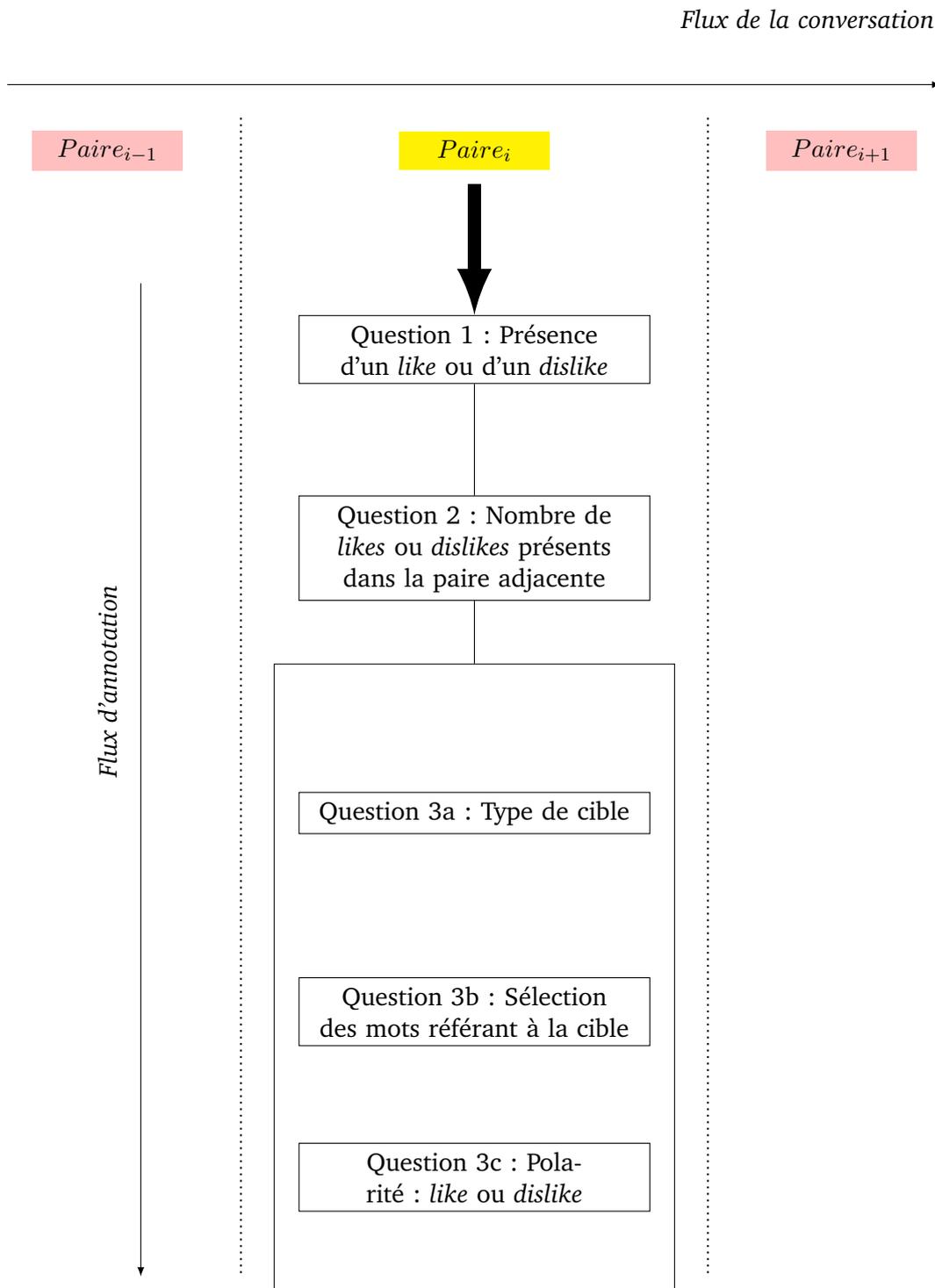


FIGURE 7.1 – Processus d'annotation de la campagne Semaine

obtenu un sous-ensemble d'annotations correspondant à un ensemble non-homogène d'annotateurs. Puisque les mesures d'accord sont conçues pour des ensembles homogènes d'annotateurs, nous avons appliqué un simple pourcentage d'accord. Même s'il semble parfois difficile pour les annotateurs de tomber d'accord sur la présence d'une expression, leur accord concernant la polarité apparaît en revanche comme plus importante. 41% des sous-ensembles montrent un accord entre 50% et 75%, et 52% des sous-ensembles ont un pourcentage d'accord supérieur à 75%. L'accord est également important concernant la cible : 61% des sous-ensembles ont un pourcentage d'accord supérieur à 50%.

### 7.1.1.3 Évaluation et discussion

À partir des 600 paires du corpus précédemment annoté, nous gardons 503 paires pour l'évaluation du système en supprimant les paires pour lesquelles un consensus ne peut être trouvé entre les 4 annotateurs - c'est-à-dire que nous gardons comme référence le vote majoritaire correspondant aux données dans lesquelles au moins trois annotateurs sont d'accord. Deux versions du système ont été testées, une version *baseline* n'intégrant pas le traitement des L&DSC (baseline-NoL&DSC) et une version intégrant le traitement de ces expressions (with-L&DSC). Nous utilisons trois mesures différentes pour évaluer les performances du système :

- détection de la présence d'un *like* ou *dislike* de l'utilisateur est évaluée par le kappa de Fleiss ;
- cohérence sur le nombre d'expressions détectées est évaluée par le coefficient alpha de Cronbach ;
- accord sur la polarité est mesuré en utilisant le kappa de Fleiss
- accord sur le type de cible avec le pourcentage d'accord.

	Baseline-NoL&DSC	with-L&DSC
Présence de L&D	$k = 0.40$	$k = 0.61$
Nb de L&D	$\alpha = 0.51$	$\alpha = 0.67$
Polarité	$k = 0.81$	$k = 0.84$

TABLEAU 7.2 – Mesures d'accord entre le système et la référence (évaluation *Small-Talk* 1)

Pour la version Baseline-NoL&DSC, concernant la détection de la présence de L&D, les performances du système montre un kappa de 0.40, ce qui est en-dessous des résultats attendus ( $< 0.6$ ). Les performances concernant le nombre de L&D par paires adjacentes est également en-dessous du résultat attendu :  $\alpha = 0.51$ . La détermination de la polarité montre en revanche de bons résultats  $k = 0.81$ .

Pour la version with-L&DSC, le nombre d'expressions détectées augmentent : nous passons de 75 à 112 *likes*, et de 22 *dislikes* à 33 *dislikes*.

L'accord entre la sortie du système et la référence est supérieur à celui obtenu avec la version baseline concernant la détection de la présence d'une expression utilisateur ( $k = 0.61$ ). Le nombre d'expressions est également mieux détecté par le système ( $\alpha$  supérieur à 0,6 contre 0.51). 4 % des paires (25 paires) sont annotées par le système comme contenant 1 expression, tandis que l'annotation référencée n'indique pas la présence d'une expression de *like* ou *dislike*. Pour 8% des paires (43 paires), c'est le phénomène inverse (1 expression annotée par la référence mais pas par le système). Le kappa obtenu pour la polarité est encourageant puisqu'il est égal à 0,84. En ce qui concerne le type de cible, nous obtenons un pourcentage d'accord de 53%. Le désaccord concerne fréquemment une confusion entre les catégories *unknown* et *other*.

Nous avons effectué une analyse approfondie du désaccord entre les sorties du système et les annotations issues de AMT. Nous avons identifié deux types principaux de difficultés.

- La première difficulté concerne le traitement de la parole spontanée. Le corpus *Semaine* contient un grand nombre d'énoncés disfluents qui perturbent la structure syntaxique du tour de parole et entravent ainsi à la fois le processus d'annotation et le système de détection. Dans la paire suivante, Agent : *'Oh !'* – Utilisateur : *"are just very good really good film and read a book"*, la structure grammaticale de la phrase utilisateur est floue (absence du sujet, présence de réparations) ce qui rend difficile l'analyse syntaxique de la phrase et donc la détection de L&D. Pour gérer cette difficulté, il serait intéressant d'intégrer un système capable d'étiqueter automatiquement les disfluences, tel que celui présenté par Dutrey et al. (2014). La structure disfluente de la phrase pourrait ainsi être intégrée à nos règles syntaxiques et sémantiques.
- La seconde difficulté concerne le manque de contexte fourni par certaines paires adjacentes. Notre système offre une première étape dans l'intégration du contexte d'interaction en considérant conjointement le tour de parole de l'utilisateur et celui de l'agent. Comme le montre l'écart entre les résultats de la version baseline et ceux de la version avec le traitement des L, cela permet d'analyser correctement un grand nombre d'expressions. Cependant, l'absence de prise en compte de paires antérieures peut provoquer des désaccords non seulement entre les sorties du système et la référence, mais aussi entre les annotateurs. Dans l'exemple suivant, Agent : *"good. ah good"* – Utilisateur : *"my favourite emotion"*, la source (ici, l'utilisateur) peut être facilement identifiée mais les informations contenues dans la paire adjacente ne sont pas suffisantes pour identifier la cible.

## 7.1.2 Évaluation qualitative *small-talk topic* : séquence thématique

### 7.1.2.1 Résultats du Niveau B

Nous testons le système sur 10 sessions du corpus *Semaine*. Au cours des 10 sessions, 23 séquences thématiques sont introduites. Pour presque toutes ces séquences, le système parvient à détecter au moins un *mot-thème* pertinent. Pour seulement 4 séquences, aucun *mot-thème* n'est détecté<sup>2</sup>. Ces séquences sont liées à 4 cadres thématiques différents : *Professional Activity*, *Project*, *Anger* and *Free-Time Activities*. Le système échoue car aucun mot correspondant au domaine lié à ces cadres thématiques ne peut être trouvé dans le tour de parole de l'utilisateur. Par exemple, l'opérateur fournit le cadre thématique *Free-Time Project* en ouvrant la séquence avec la phrase : *"tell me more about your plans for the weekend"*. Cependant, le seul *mot-thème* pertinent prononcé par l'utilisateur, *"gardenening"*, appartient à un domaine que nous n'avons pas lié au cadre thématique courant, car il n'apparaît pas pertinent pour la plupart des séquences.

Ces résultats sont assez encourageants concernant le silence, cependant ils pourraient être améliorés en ce qui concerne le bruit. En effet, la plupart des erreurs sont dues à des erreurs de repérage de *POS-tagging* ou à la trop large couverture de certains domaines. Lorsqu'un domaine couvre un grand nombre de *synsets*, le système peut récupérer des mots qui ne sont pas liés au thème courant.

### 7.1.2.2 Fusion des résultats (niveau C)

En ce qui concerne la phase de fusion (niveau C1), la relation sémantique la plus fréquente que le système établit est la relation de similarité. Le résultat est sensiblement le même pour l'identification de relation entre cible et thème (niveau C2). Aucun synonyme entre *cible* et *mot-thème* n'a été trouvé dans les 10 sessions. Afin de valider la pertinence de cette relation, il est nécessaire de tester le système sur un plus grand nombre de sessions. Un grand nombre de *likes*

2. Ces séquences font partie des sessions 66, 90 et 114.

et de *dislikes* correspond à une cible qui n'est pas identifiée (environ 40 %) au niveau C1. Au niveau C3, notre système montre son efficacité en les associant au sujet actuel.

## 7.2 Évaluation dans le cadre d'un scénario de négociation

Afin de pouvoir constituer un ensemble d'annotation de référence, nous avons mené une campagne d'annotations. 77 sessions transcrites et extraites du corpus (DeVault et al., 2015) ont été utilisées.

### 7.2.1 Protocole et plateforme d'annotation

Les annotateurs ont été recrutés via la plateforme *CrowdFlower*. Ils ont été sélectionnés en fonction du taux d'approbation qui leur a été attribué selon la qualité de leur travail. Chaque session a été annotée par 4 annotateurs différents et chaque annotateur n'a annoté qu'une seule session. L'attribution d'un annotateur à une session particulière a été décidée au hasard. Une fois les annotateurs recrutés, ils sont automatiquement redirigés vers une plateforme d'annotation externe, où ils doivent répondre à une série de questions sur la présence de L&D de l'utilisateur ciblant les objets impliqués dans la négociation (un tableau, un ensemble de vinyles, deux chaises, une horloge, un ensemble d'assiettes et deux lampes).

À la manière de celle développée par l'évaluation Semaine, cette plateforme définit la paire adjacente comme le contexte d'annotation. Néanmoins, elle a été conçue pour donner davantage de contexte aux annotateurs en mettant en valeur la progression de la conversation<sup>3</sup>. Ainsi, à la différence de ce que nous avons proposé avec la plateforme précédente, notre volonté a été de ne pas soustraire le contexte antérieur de la paire annotée : l'utilisateur devait pouvoir visualiser, pour toute nouvelle paire, les paires précédemment annotées. Pour cette raison, la page d'annotation a été pensée de manière ergonomique : une division en 3 parties, avec à gauche une zone indépendante de la zone d'annotation proprement dite, où l'annotateur peut remonter librement le cours de la conversation et qui contient les précédentes paires adjacentes mais également les annotations associées. La zone d'annotation, à droite de la page, a quant à elle été divisée en deux espaces : le premier mettant en valeur la paire courante à annoter, et en dessous, deux questionnaires à remplir. Que ce soit pour la première ou la seconde tâche, l'intérêt était d'interroger sur les mêmes objets à chaque paire adjacente afin de pouvoir évaluer les changements d'avis de l'utilisateur au cours de la négociation.

### 7.2.2 Accord inter-annotateur

	Like		Dislike	
	Kappa	Accord	Kappa	Accord
Painting	0.35	79.8	0.20	85.2
Lamp	0.45	66.2	0.19	93.05
Record	0.35	65.1	0.30	97.7
Chair	0.46	61.6	0.10	87.4
Plate	0.45	69.8	0.31	82.2
Clock	0.36	85.1	0.30	87.7

TABLEAU 7.3 – Moyennes des mesures obtenues pour chaque objet

3. La plateforme est décrite en détail en Annexe D

Pour calculer l'accord inter-annotateur, nous avons utilisé le Kappa de Fleiss (Fleiss, 1971). Tel que le schéma d'annotation est défini, l'annotateur est en mesure de sélectionner plusieurs objets pour chacune des questions posées et au sein d'une même paire adjacente, un objet peut être identifié comme la cible d'un *like* et d'un *dislike*. Par exemple, l'annotateur peut identifier une expression de *like* et une expression de *dislike* envers la lampe, une expression de *like* concernant les disques et un *dislike* concernant la peinture. Afin de pouvoir prendre en charge cette problématique de *multi-label* dans le cadre de la mesure de l'accord inter-annotateurs, nous nous sommes appuyée sur la méthode décrite par Core and Allen (1997) et Di Eugenio et al. (1998) et avons calculé les kappas sur la base de cette structure hiérarchique des annotations. Pour chaque session et pour chaque objet, nous avons mesuré le Kappa de Fleiss et le pourcentage d'accord pour, d'une part, les *likes*, et d'autre part, les *dislikes* (voir le tableau 7.3). Les scores obtenus montrent une certaine variabilité qui peut être expliquée par la nature même de la tâche d'annotation. Il est en effet possible que l'analyse à grain fin demandée aux annotateurs – identification d'expressions de *likes* et de *dislikes* pour des objets différents – ait pu favoriser le désaccord des quatre annotateurs. La difficulté a pu également être favorisée par la nature du phénomène linguistique. En raison de la variabilité de leurs expressions, l'identification conjointe des *likes* et des *dislikes* et de leurs cibles peut être difficile. Certaines expressions sont plus complexes et reposent sur des indices linguistiques marqués sur le plan axiologique et affectif. Comme expliqué par Marcheggiani et al. (2014), ce type d'expressions peut provoquer des scores de désaccord plus élevés. Cette hypothèse semble être confirmée par le fait que les scores les plus faibles concernent les *dislikes*. Comme l'a révélé une analyse approfondie des annotations, les expressions des *likes* sont plus marquées que celles des *dislikes*. L'utilisateur emploie des indices linguistiques plus ambigus pour exprimer une aversion que pour un goût, ce qui peut rendre plus difficile leur identification par les annotateurs.

### 7.2.3 Évaluation et discussion

La référence utilisée pour l'évaluation du système a été construite sur la base du consensus trouvé entre les annotateurs. Est définie comme un consensus, toute annotation fournie par la majorité (3 sur 4 annotateurs). Pour chaque  $Object_o$ , chaque paire adjacente  $AP_p$ , nous avons défini la valeur de  $Like_{Object_o}(AP_p) = \{True, False\}$  et  $Dislike_{Object_o}(AP_p) = \{True, False\}$  en fonction de la valeur donnée par au moins trois annotateurs. Si, pour  $Like_{Object_o}(AP_p)$  ou  $Dislike_{Object_o}(AP_p)$ , aucun consensus n'est trouvé, l'annotation n'est pas ajoutée à la référence. 12790  $Like_{Object_o}$  et  $Dislike_{Object_o}$  ont été conservés sur les 13326 éléments initiaux. A la manière de l'évaluation *Small-Talk 1*, deux versions du système ont été testées : une version baseline, *Baseline-Nego*, ne prenant pas en compte les L&DSC, et une version étoffée prenant en compte ces expressions, *Nego-WithL&DSC*.

Nous utilisons le kappa de Fleiss pour évaluer l'accord entre la référence et le système. Tout d'abord, pour un aperçu des performances du système, nous mesurons les kappas pour les 12790 éléments, c'est-à-dire sans distinction d'objet ni de polarité. Ensuite, nous sélectionnons les éléments en fonction du type d'expression : *like* ou *dislike*. Deux groupes sont créés : l'un rassemblant toutes les annotations concernant les *likes*, l'autre rassemblant toutes les annotations concernant les *dislikes*. Enfin, nous effectuons une discrimination sur le type d'objet et le type d'expression.

La version *Baseline-Nego* montre des résultats inférieurs à ceux de la version *Nego-WithL&DSC*. Les mesures d'accord sur l'ensemble des annotations (*likes* et *dislikes*, et toutes cibles confondues) donnent un score de  $k = 0.51$ , pour la version baseline, contre  $k = 0.61$ , pour la version *Nego-WithL&DSC* (voir le tableau 7.4). Ce dernier résultat est substantiel (George and Mallery, 2010). En ce qui concerne la distinction de polarité, i.e la distinction entre *likes* et *dislikes*, la version baseline montre, pour les *likes*, un kappa de 0.51 et, pour les *dislikes*, un kappa de 0.43. Ces résultats sont supérieurs pour la version *Nego-WithL&DSC* :  $k = 0.62$ , pour les *likes*, et  $k = 0.53$ , pour les *dislikes*. Les mesures faites avec une distinction sur les types de cibles montrent des

résultats hétérogènes. Toutefois, ceux de la version *Nego-WithL&DSC* sont ici aussi supérieurs à ceux de la version baseline.

	<i>Baseline-Nego</i>		<i>Nego-WithL&amp;DSC</i>	
	Like	Dislike	Like	Dislike
All Objects	0.51	0.43	0.62	0.53
Painting	0.41	0.52	0.54	0.61
Lamp	0.55	0.49	0.67	0.60
Record	0.44	0.52	0.56	0.66
Clock	0.54	0.60	0.67	0.73
Chair	0.49	0.59	0.61	0.77
Plate	0.41	0.52	0.43	0.60

TABLEAU 7.4 – Mesures d'accord entre le système et la référence (évaluation négociation)

Concernant les résultats cible par cible, une analyse qualitative des résultats de la version *Nego-WithL&DSC* montre que les résultats inférieurs ne sont dus ni à la nature de la cible ni à la défaillance du système dans l'identification de ce type de cible. Trois principaux facteurs de désaccord apparaissent. Tout d'abord, les annotateurs montrent une compétence naturelle pour lire des implicatures conversationnelles (Grice, 1975) et pour inférer des stratégies de négociation chez l'utilisateur. Par exemple, de nombreuses questions interrogeant l'agent sur ses préférences, comme "Do you like the painting", ont été annotées comme l'indice d'un *like* de l'utilisateur. Nous avons délibérément choisi de ne pas traiter de telles implicatures. Leur lecture suppose en effet d'attribuer arbitrairement un sens à une phrase dénotant autre chose, sans aucune certitude sur le sens sous-jacent attribué. Pour éviter les erreurs d'interprétation et les mauvais résultats, le modèle ne traite que des indices linguistiques objectifs tels que des éléments lexicaux ou les structures syntaxiques. Ensuite, certaines caractéristiques linguistiques distinctives peuvent expliquer les silences produits par le système. Certains *likes* ou *dislikes* sont exprimés à travers des structures de phrases non traitées par les grammaires du fait de leur particularisme. Enfin, la variabilité du discours conversationnel peut perturber l'analyse. L'extrait suivant en donne un aperçu :

- (1a) Agent : "What do you think about me getting one record and both lamps and you getting two records and the painting?"  
 (1b) User : "Hmm so both records so I'll take two"  
 (2a) Agent : "Yeah exactly"  
 (2b) User : "And the lamp I'm sorry and the painting"

Dans ces deux paires, le système ne parvient pas à détecter les deux goûts envers la peinture et les disques. D'abord, il détecte une question dans le tour de parole de l'agent (1a), mais ne trouve aucune marque d'approbation pertinente dans le tour de parole de l'utilisateur (1b). Ensuite, dans le tour de parole de l'utilisateur (1b), du fait de la position du complément d'objet, dans la phrase "Hmm so both records so I'll take two", l'expression de *like* ne peut être détectée. Enfin, le manque de contexte fourni par la seconde paire adjacente et l'autocorrection faite par l'utilisateur, invalident la possibilité de relier la peinture au *like* exprimé dans la première paire.

## 7.3 Conclusion du chapitre 7

Dans ce chapitre, nous avons présenté trois évaluations. Deux d'entre elles ont été menées sur le corpus Semaine, et le troisième sur le corpus de négociation. Deux de ces évaluations ont été pour nous l'occasion de vérifier les performances du système pour la détection des *likes* et *dislikes* au niveau du tour de parole et de la paire adjacente, dans deux cadres applicatifs différents, négociation et *small-talk*. Ces évaluations de nature quantitative ont nécessité la création de jeux d'annotation. Dans les deux cas, les annotateurs ont été recrutés via des plateformes de *crowdsourcing* : *Amazon Mechanical Turk*, pour l'évaluation sur le corpus Semaine, et *Crowd-Flower*, pour l'évaluation sur le corpus de négociation. La collecte des annotations s'est ensuite déroulée grâce à deux plateformes web que nous avons nous-mêmes développées. Ces deux plateformes supportent des schémas d'annotation partageant des caractéristiques communes : la mise en valeur du contenu verbal, la prise en compte des caractéristiques conversationnelles et la simplification de la tâche d'annotation. À partir des annotations obtenues pour chacune des deux tâches d'annotation, nous avons constitué deux références pour évaluer notre système. Dans le cadre de ces deux évaluations, le système a été testé sous deux versions : une *baseline*, n'incluant pas le traitement des expressions de *likes* et *dislikes* strictement collaboratifs (L&DSC). Pour les deux évaluations, les résultats.

- **Small-Talk 1** : l'évaluation sur le corpus Semaine montrent des résultats intéressants. Pour la version baseline, nous avons un kappa de 0.40 concernant la présence d'une expressions de *like* ou *dislike*, et un alpha de Cronbach de 0.51, pour la détermination du nombre d'expressions présentes. La version étoffée obtient de meilleurs résultats :  $k = 0.61$  pour la présence d'un *like* ou d'un *dislike*, et  $\alpha = 0.67$ . Ces derniers résultats sont encourageant et substantiels au regard des seuils établis. De plus, la détection de la polarité obtient un kappa de 0.81.
- **Négociation** : les mesures d'accord entre le système et la référence montrent également un écart entre la version baseline et la version étoffée. La version baseline obtient un kappa de 0.51 pour la détection des *likes* sur l'ensemble des objets de la négociation, et un kappa de 0.43 pour la détection des *dislikes* toujours sur l'ensemble des objets de la négociation. Les résultats de la version étoffée prenant en compte les expressions strictement collaboratives sont supérieurs :  $k = 0.62$ , pour la détection des *likes*, sans distinction sur les objets de la négociation, et  $k = 0.53$ , pour les *dislikes*, toujours sans distinction sur les objets de la négociation.

Dans ces deux évaluations, l'écart entre les résultats des versions baseline et des versions incluant des *likes* et des *dislikes* strictement collaboratifs montre que la prise en compte de ces expressions est nécessaire pour améliorer la qualité de la détection.

La troisième évaluation, menée sur le corpus Semaine, est une évaluation qualitative visant à vérifier le bon fonctionnement des analyses menées dans le cadre de la séquence thématique. Dans ce cadre, le système a été testé sur dix sessions du corpus. Les résultats montrent une bonne détection de mots-thèmes. La phase de fusion apparaît comme s'effectuant correctement. Néanmoins, pour valider ces résultats, il sera nécessaire de procéder à une évaluation quantitative comparant les résultats du système à ceux d'annotateurs humains et exploitant un corpus plus large.



# Chapitre 8

## Conclusion et perspectives

### Sommaire

---

8.1	Résumé . . . . .	<b>108</b>
8.2	Contributions . . . . .	<b>109</b>
8.2.1	Formalisation des expressions d'attitude en contexte conversationnel . . . . .	109
8.2.2	Implémentation et validation des règles linguistiques . . . . .	109
8.2.3	Méthodologie d'évaluation . . . . .	110
8.3	Perspectives . . . . .	<b>110</b>
8.3.1	Amélioration des performances du système . . . . .	110
8.3.2	Évaluation dans le cadre d'une interaction humain-agent . . . . .	111
8.3.3	Vers la gestion des préférences de l'utilisateur . . . . .	111
8.3.4	Vers l'hybridation . . . . .	112

---

L'objectif de notre thèse est le développement d'un modèle de détection des attitudes exprimées par l'utilisateur au cours de conversations humain-agent. Notre travail se situe à la croisée de deux domaines, celui des agents conversationnels animés et celui de l'analyse de sentiments. En l'absence de travaux antérieurs se concentrant sur ces phénomènes tels qu'ils s'expriment dans le contenu verbal de l'utilisateur, nous avons cherché à développer un modèle de détection s'inspirant des avancées menées dans le domaine de l'analyse de sentiment sur des textes écrits, qui puisse être adapté au contexte spécifique des conversations humain-agent. Dans ce mémoire, nous avons présenté les différents aspects nécessaires au développement d'un tel modèle. Se fondant sur une approche linguistique du problème, tant dans un objectif de formalisation que d'implémentation, il place au centre l'analyse du contenu verbal. En observant et formalisant le phénomène tel qu'il se réalise dans des conversations humain-agent, nous avons produit un ensemble de règles sémantiques et syntaxiques. Si le système a été pensé pour gérer les caractéristiques propres à ce contexte, les différentes évaluations que nous avons menées ont également eu toutes – à des degrés différents – pour but de mettre en relief la manière dont le contenu sémantique des expressions d'attitude est ancré à l'intérieur d'unités conversationnelles telles que le tour de parole, la paire adjacente ou la séquence thématique. Afin de pouvoir vérifier le niveau de généralité des règles, deux différents cadres d'analyse ont été pris en compte : celui de la négociation et celui des conversations de type *small-talk*. Cette conclusion est pour nous l'occasion de dresser le bilan de cette approche centrée sur l'analyse linguistique, dans un premier temps, en résumant les différentes étapes de notre travail puis, dans un deuxième temps, en rappelant nos différentes contributions, et enfin, dans un troisième temps, en proposant des perspectives à notre travail.

## 8.1 Résumé

Cette section reprend et synthétise les différents points abordés dans chacun des chapitres composant de ce mémoire.

- **Partie I** : Dans un premier temps, nous avons présenté une vue d'ensemble des principaux axes de recherche en analyse de sentiments. Nous avons ainsi fait ressortir le fait que, parallèlement à des travaux de classification, des études proposent une analyse détaillée des expressions de sentiments. Leur approche repose sur une exploration des structures sémantiques et syntaxiques pour déterminer avec plus de précision la valeur des attributs *polarité* et *cible*. Forte des données rassemblées au cours de l'état de l'art, nous avons été en mesure de définir le type de méthode qui nous semblait appropriée.
- **Partie II** : Dans un deuxième temps, nous nous sommes attachée à décrire les fondements linguistiques de la notion d'attitude, en étudiant leurs mécanismes sémantiques et syntaxiques à différents niveaux : celui du tour de parole, celui de la paire adjacente et celui de la séquence thématique.
  - Dans le cadre du tour de parole (Chapitre 3), sur la base des patrons syntaxiques proposés dans la littérature, nous avons analysé les structures de phrase et les mécanismes sémantiques par lesquels la source, la cible et la polarité sont définis. Nous avons proposé une optimisation et une extension des patrons déjà existant, afin de regrouper au sein d'une même structure, pour chaque expression, la question de l'attribution des rôles sémantiques de source et de cible, les procédés sémantiques de catégorisation de l'expression et les règles sémantiques de calcul de polarité.
  - Dans le cadre de la paire adjacente (Chapitre 4), nous avons pu mettre en valeur la nature participative et contributive des expressions d'attitude – en partie déterminées par le contenu verbal qui les précèdent. Ce caractère participatif est particulièrement saillant dans le cas d'attitudes que nous avons nommé *strictement collaborative* (EASC), dont l'interprétation repose exclusivement sur la prise en compte du contenu sémantique des énoncés de l'agent et de l'utilisateur. Elles se distinguent des expres-

sions simples d'attitude étudiées dans le chapitre 3. Nous avons étudié leur fonctionnement et en avons proposé une représentation formelle exploitable dans un modèle computationnel.

- Dans le cadre de la séquence thématique (Chapitre 5), nous avons analysé la manière dont les cibles des attitudes sont liées au thème courant de la conversation.
- **Partie III** : Dans le premier volet de la troisième partie (chapitre 6), nous avons décrit l'architecture de notre système. Celui-ci se compose de trois principaux traitements, dédiés pour chacun à l'un des trois cadres d'analyse décrits dans la partie II. Ce système se concentre sur un sous-phénomène particulier, celui des *likes* et des *dislikes* que nous avons considéré comme une sous-catégorie transverse à celle des attitudes. Dans le second volet (Chapitre 7) de cette partie, nous avons enfin présenté les trois différentes évaluations que nous avons menées pour évaluer le système. Chacune d'elles est consacrée à un sous-aspect spécifique décrit dans le chapitre 6 : la détection des *likes* et *dislikes* au sein de la paire de l'énoncé et de la paire adjacente, et l'analyse sémantique des cibles au sein de la séquence thématique. Deux évaluations sont consacrées au tour de parole et à la paire adjacente : pour la première nous avons exploité 600 paires extraites du corpus Semaine, pour la seconde nous avons exploité 77 sessions du corpus de négociation de Gratch et al. (2016). Pour chacune de ces deux évaluations, nous avons constitué un ensemble d'annotations produites par des annotateurs recrutés sur des plateformes de *crowdsourcing* : *Amazon Mechanical Turk*, pour la première et *Crowdfunder*, pour la seconde.

## 8.2 Contributions

### 8.2.1 Formalisation des expressions d'attitude en contexte conversationnel

Notre travail nous a permis d'analyser en profondeur les attitudes telles qu'elles s'expriment dans le contenu verbal de l'utilisateur au cours d'une conversation avec un agent conversationnel animé. Au terme de cette analyse, nous avons été en mesure de proposer une formalisation des attitudes prenant en compte trois cadres différents : le tour de parole, la paire adjacente et la séquence thématique. Dans le tour de parole, nous avons ainsi rassemblé les données éparses de la littérature et proposé une formalisation syntaxique et sémantique des énoncés référant à une attitude. L'ensemble de patrons qui résultent de ce travail offrent une représentation des mécanismes de compositionnalité définitoires des valeurs des attributs *source*, *cible* et *polarité*. Dans la paire adjacente, nous avons pu proposer une formalisation des mécanismes sémantiques se jouant entre les deux tours de parole. Enfin, dans le cadre de la séquence thématique, nous avons proposé une analyse des relations sémantique existant entre les cibles et le thème courant.

### 8.2.2 Implémentation et validation des règles linguistiques

À l'issue des analyses présentées dans la partie II, un ensemble de règles syntaxiques et sémantiques a été développé. Ces règles prennent en charge différents niveaux de granularité - syntagme, phrase, tour de parole, paire adjacente et séquence thématique. Deux types de règles ont été créées : des règles sémantiques et des règles logico-syntaxiques, combinant analyse syntaxique et analyse de la structure actancielle des prédicats.

- Les règles logico-sémantiques sont appliquées aux niveaux syntagmatique et phrastique pour la détection des *expressions simples de likes et dislikes* dans le tour de parole de l'utilisateur. Elles sont implémentées sous forme de grammaires via la plateforme *Unitex*. Les grammaires prennent en charge la reconnaissance de plusieurs patrons syntaxiques et intègrent la reconnaissance des rôles sémantiques *source* et *cible* ainsi que des règles de calcul de la polarité.

- Les règles sémantiques sont appliquées au niveau de la paire adjacente et de la séquence thématique. Au niveau de la paire adjacente, elles permettent le calcul de la polarité et l'attribution des rôles sémantiques pour les *expressions d'attitude strictement collaboratives*. Au niveau de la séquence thématique, elle permettent d'établir les relations sémantiques pouvant exister entre la cible et le thème courant de la conversation.

Les deux évaluations quantitatives menées respectivement sur les corpus négociation et *Semaine* ont montré des résultats encourageants.

- L'évaluation réalisée avec les annotations collectées sur le corpus *Semaine* met en évidence la pertinence d'un traitement des *L&DSC*. Concernant la présence d'une expression dans la paire adjacente courante, la version *baseline* du système – ne traitant pas les *L&DSC* – obtient un kappa de 0.40, tandis que la version prenant en compte les *L&DSC* obtient un kappa de 0.61. Concernant le nombre d'expressions détectées, là encore les scores de la version prenant en compte les *L&DSC* ( $\alpha = 0.67$ ) sont meilleures que ceux de la version *baseline* ( $\alpha = 0.51$ ).
- L'évaluation réalisée dans le cadre d'une application en contexte de négociation montre également une différence entre la version *baseline* et la version intégrant les expressions strictement collaboratives. Les résultats pour chacun des objets de la négociation montrent des kappas inférieurs de 0.10 à 0.20 pour la version *baseline* par rapport à la version étoffée.

### 8.2.3 Méthodologie d'évaluation

Afin d'évaluer notre système, nous avons cherché à développer une méthodologie d'évaluation permettant de prendre en compte les caractéristiques conversationnelles du contexte d'application. Pour cela, nous avons développé deux plateformes d'annotation mettant en valeur la paire adjacente comme contexte d'annotation. De plus, la plateforme développée dans le cadre de l'évaluation négociation offre un contexte d'annotation concentré sur la paire adjacente, mais fait également émerger la progression de la conversation en permettant à l'annotateur de conserver un accès à l'historique des tours de parole et à ses précédentes annotations. Du plus, ces plateformes ont été conçues pour faciliter la tâche d'annotation en proposant à l'annotateur à un ensemble des questionnaires relatifs à la présence d'expressions et à leurs valeurs sémantiques.

## 8.3 Perspectives

### 8.3.1 Amélioration des performances du système

Les différentes évaluations menées au cours de notre thèse ont mis en évidence un certain nombre de points à améliorer dans notre système.

1. Au niveau du tour de parole, la détection pourra être améliorée par une meilleure prise en charge des disfluences. Comme expliqué dans le chapitre 3, les grammaires intègrent une certaine flexibilité pour la détection des attitudes exprimées oralement. Néanmoins, les irrégularités syntaxiques des énoncés de nos corpus se sont avérées parfois très complexes. Les évaluations menées sur le corpus de négociation et le corpus *Semaine* montrent des exemples pour lesquels le système n'a pas été en mesure – malgré la flexibilité que nous y avons apportée – de gérer des énoncés très disfluents. Une première étape pourra intégrer un moteur de détection des disfluences comme celui développé par Dutrey (2014), corrigeant ces cas complexes, lesquels pourront ensuite être traités par notre système de détection.
2. Au niveau de la paire adjacente, l'une des pistes d'amélioration est la prise en charge des questions ouvertes dans le traitement de la collaboration. Dans le chapitre 4, nous

avons proposé des règles pour déterminer la valeur sémantique d'expressions strictement collaboratives déclenchées à partir de questions ouvertes posées par l'agent. Néanmoins, celles-ci n'ont pas été implémentées dans la version actuelle du système. Ces expressions sont fréquentes notamment dans le corpus de négociation. L'intégration de leur traitement dans le système pourrait donc améliorer le taux de rappel.

3. Au niveau de la séquence thématique, l'analyse des relations sémantiques entre les cibles et le thème courant concerne la synonymie et la similarité sémantique. Dans le chapitre 5, nous avons également identifié des relations de méronymie/holonymie et d'hyponymie/hyperonymie. L'intégration de ces relations dans l'analyse permettrait un ancrage plus poussé dans la structure thématique.

### 8.3.2 Évaluation dans le cadre d'une interaction humain-agent

Au cours de notre thèse, nous avons pu mener trois évaluations permettant de tester l'efficacité de nos règles et de nos grammaires. Ces évaluations ont été réalisées sur des transcriptions de conversations ayant déjà eu lieu. Cette étape, qui s'est révélée nécessaire dans notre travail, devra néanmoins être complétée dans de futurs travaux par une évaluation au cours d'une interaction entre un utilisateur humain et un agent conversationnel. Pour cela, il sera nécessaire d'intégrer notre système dans une plateforme d'agent virtuel, comme la plateforme Greta (de Sevin et al., 2010), par exemple. Ce type d'évaluation aura pour but, premièrement, de tester les performances des règles sur un dialogue en temps réel transcrit par un moteur de synthèse vocale. L'objectif sera de déterminer si les règles ne produisent pas plus de bruit et de silence dans ce cadre et si le temps de traitement est suffisant pour suivre de rythme du *turn-taking*. Dans un second temps, l'évaluation permettra de valider la pertinence des sorties de notre système pour une exploitation par un moteur de dialogue ou un planificateur d'intention de l'agent (Heylen et al., 2008) et d'approuver leur utilité pour le développement de stratégies de dialogue et d'interaction. Cette évaluation pourra être réalisée dans les deux types d'interactions envisagées dans notre thèse : le *small-talk* et la négociation.

### 8.3.3 Vers la gestion des préférences de l'utilisateur

L'une des perspectives de notre travail est l'évolution vers un système de détection des préférences de l'utilisateur. À la différence des attitudes et des *likes* et *dislikes*, nous considérons les préférences comme étant ordonnées les unes par rapport aux autres. Un système de détection des préférences doit être capable de fournir un modèle utilisateur reproduisant cet ordre. Dans le cadre de la négociation par exemple, les résultats doivent permettre d'identifier l'objet se plaçant en première position dans l'ordre des préférences de l'utilisateur et celui occupant la dernière position.

La connaissance des préférences de l'utilisateur telles qu'elles s'expriment dans le contenu verbal peut s'avérer une source d'information utile dans le cadre d'application telle que la négociation. Comme l'expliquent Ouali et al. (2018), « la collaboration entre interlocuteurs pour la satisfaction de buts implique donc la confrontation des préférences et des expertises de chacun des partis, ce qui les conduit à négocier sur la manière de satisfaire le but en prenant en compte les préférences des deux partis ». Une extraction des préférences de l'utilisateur offrirait ainsi un support à l'agent pour le développement de ses stratégies de négociation.

**Expressions verbales des préférences** Dans le contenu verbal, les expressions comparatives et superlatives sont des formes dédiées à l'expression de préférences. Ces deux types de structures mettent en valeur des objets les uns par rapport aux autres. Concernant les structures comparatives, elles mettent en relation deux cibles entre lesquelles est posée une évaluation comparative de leurs propriétés respectives. Quant aux structures superlatives, dans leur forme

relative, elles permettent de faire émerger au sein d'un groupe une cible mise en valeur pour sa supériorité ou son infériorité.

Néanmoins, les expressions de *likes* et de *dislikes* détectées dans la version actuelle du système servent également à l'expression de préférences. À la différence des structures comparatives et superlatives, la mise en relation des différents objets – cibles – ne se fait pas à l'intérieur d'une même structure phrastique, mais peut avoir lieu entre différentes phrases voire entre les différents tours de parole.

**Détecter les expressions verbales de préférences** Afin de faire évoluer le système vers le traitement des préférences de l'utilisateur, certaines règles doivent être ajoutées aux règles existantes. Dans un premier temps, notre objectif est de développer des grammaires permettant de traiter les structures comparatives et superlatives décrites ci-dessus. Dans un second temps, notre travail devra se concentrer sur la manière dont les règles actuelles peuvent être exploitées pour la création d'un modèle des préférences de l'utilisateur. L'une des pistes est la mise en relation des résultats de la paire courante avec l'historique des résultats antérieurs. Tout *like* ou *dislike* détecté devra être comparé aux *likes* et *dislikes* précédents afin de vérifier la supériorité ou l'infériorité d'une cible sur les autres.

**Annoter les Préférences de l'Utilisateur** La création d'un ensemble d'annotations de préférences de l'utilisateur est une étape indispensable à la création d'un modèle de détection. Pour cela, une plateforme d'annotation a déjà été développée (voir l'annexe D). Son architecture reprend en partie le modèle de la plateforme décrite dans le chapitre 7. Le cadre d'annotation est celui de la paire adjacente, néanmoins, la plateforme met en valeur la progression de la conversation en donnant à l'annotateur un accès à l'historique des paires adjacentes déjà traitées. L'annotation se fait par l'intermédiaire de questionnaires. Le modèle d'annotation sur lequel repose la plateforme prend en compte un ensemble limité d'objets. L'objectif de l'annotation est d'établir un classement de ces objets en fonction de préférences exprimées par l'utilisateur. Pour chaque paire adjacente, l'annotateur attribue un score à chacun des objets. Chaque score indique la position de l'objet dans l'échelle des préférences de l'utilisateur.

Par exemple, lorsque l'utilisateur exprime verbalement une préférence pour un tableau sur des disques et des lampes, l'annotateur peut attribuer au tableau le score 3 et aux lampes et aux disques, le score 1. À chaque nouvelle paire, l'annotateur note ainsi l'évolution du modèle de préférences, en fonction de nouvelles données fournies par l'utilisateur.

### 8.3.4 Vers l'hybridation

À plus long terme, une ouverture nous semble particulièrement fondamentale : l'intégration des règles développées dans un système hybride exploitant un algorithme d'apprentissage automatique. Ce type de système présente l'intérêt de disposer, à la manière de méthodes à base de règles, d'une modélisation détaillée des expressions de sentiments et, à la manière des méthodes d'apprentissage automatique, d'une plus grande flexibilité d'utilisation et d'une capacité de généralisation. Notre thèse ayant déjà formalisé le phénomène et testé la validité des règles produites, le travail doit se concentrer sur la création d'un modèle intégrant les descripteurs les plus pertinents.

Les travaux de Valentin Barrière œuvrent dans ce sens. Dans son article (Barriere, 2017), il présente différents modèles de HCRF (*hidden conditional random fields*) pour la détection des attitudes de l'utilisateur dans une conversation avec un agent. Les modèles testés prennent en compte différents descripteurs : des représentations distribuées de mots et des descripteurs linguistiques lexico-syntaxiques. Le modèle a été appris en partie sur le jeu d'annotations que nous avons rassemblé sur le corpus Semaine.

# Bibliographie

- Adam, C., Cavedon, L., and Padgham, L. (2010). Flexible conversation management in an engaging virtual character. In *International Workshop on Interacting with ECAs as Virtual Characters*, page 1.
- Baccianella, S., Esuli, A., and Sebastiani, F. (2010). Sentiwordnet 3.0 : an enhanced lexical resource for sentiment analysis and opinion mining. In *LREC*, volume 10, pages 2200–2204.
- Barriere, V. (2017). Hybrid models for opinion analysis in speech interactions. In *Proceedings of the 19th ACM International Conference on Multimodal Interaction*, pages 647–651. ACM.
- Beale, R. and Creed, C. (2009). Affective interaction : How emotional agents affect users. *International journal of human-computer studies*, 67(9) :755–776.
- Bednarek, M. (2009). Language patterns and attitudes. *Functions of Language*, pages 165–192.
- Bloom, K., Garg, N., and Argamon, S. (2007). Extracting appraisal expressions. *HLT-NAACL*, pages 165–192.
- Brody, S. and Elhadad, N. (2010). An unsupervised aspect-sentiment model for online reviews. In *Human Language Technologies : The 2010 Annual Conference of the North American Chapter of the Association for Computational Linguistics*, pages 804–812. Association for Computational Linguistics.
- Brown, G. and Yule, G. (1983). *Discourse analysis*. Cambridge university press.
- Button, G. and Casey, N. (1985). Topic nomination and topic pursuit. *Human studies*, 8(1) :3–55.
- Chollet, M., Ochs, M., and Pelachaud, C. (2014). From non-verbal signals sequence mining to bayesian networks for interpersonal attitudes expression. In *International conference on intelligent virtual agents*, pages 120–133. Springer.
- Church, K. W. and Hanks, P. (1990). Word association norms, mutual information, and lexicography. *Computational linguistics*, 16(1) :22–29.
- Clark, H. H. and Schaefer, E. F. (1989). Contributing to discourse. *Cognitive Science*, 13 :259–294.
- Clavel, C. and Callejas, Z. (2016). Sentiment analysis : from opinion mining to human-agent interaction. *IEEE Transactions on affective computing*, 7(1) :74–93.
- Clore, G. L., Ortony, A., and Foss, M. A. (1987). The psychological foundations of the affective lexicon. *Journal of personality and social psychology*, 53(4) :751.
- Core, M. G. and Allen, J. (1997). Coding dialogs with the damsl annotation scheme. In *AAAI fall symposium on communicative action in humans and machines*, volume 56. Boston, MA.

- Cronbach, L. (1951). Coefficient alpha and the internal structure of tests. *Psychometrika*, 16(3) :297–334.
- De Rosis, F., Pelachaud, C., Poggi, I., Carofiglio, V., and De Carolis, B. (2003). From greta’s mind to her face : modelling the dynamics of affective states in a conversational embodied agent. *International journal of human-computer studies*, 59(1-2) :81–118.
- de Sevin, E., Niewiadomski, R., Bevacqua, E., Pez, A.-M., Mancini, M., and Pelachaud, C. (2010). Greta, une plateforme d’agent conversationnel expressif et interactif. *Technique et science informatiques*, 29(7) :751.
- DeVault, D., Mell, J., and Gratch, J. (2015). Toward natural turn-taking in a virtual human negotiation agent. In *AAAI Spring Symposium on Turn-taking and Coordination in Human-Machine Interaction*. AAAI Press, Stanford, CA.
- Di Eugenio, B., Moore, J. D., Jordan, P. W., and Thomason, R. H. (1998). An empirical investigation of proposals in collaborative dialogues. In *Proceedings of the 36th Annual Meeting of the Association for Computational Linguistics and 17th International Conference on Computational Linguistics-Volume 1*, pages 325–329. Association for Computational Linguistics.
- Ding, Y., Pelachaud, C., and Artieres, T. (2013). Modeling multimodal behaviors from speech prosody. In *International Workshop on Intelligent Virtual Agents*, pages 217–228. Springer.
- Dittmar, N. (1992). Introduction : Topic in grammar and discourse. *Linguistics*, 30(1) :1–4.
- Downing, A. (2000). Talking topically. *CIRCLE of Linguistics Applied to Communication (CLAC)*, 3 :31–50.
- Dutrey, C. (2014). *Disfluency analysis and automatic detection in conversational spontaneous speech*. Theses, Université Paris Sud - Paris XI.
- Dutrey, C., Clavel, C., Rosset, S., Vasilescu, I., and Adda-Decker, M. (2014). A crf-based approach to automatic disfluency detection in a french call-centre corpus. In *Interspeech*, page to appear.
- Dzogang, F., Lesot, M.-J., Rifqi, M., and Bouchon-Meunier, B. (2010). Expressions of graduality for sentiments analysis—a survey. In *Fuzzy Systems (FUZZ), 2010 IEEE International Conference on*, pages 1–7. IEEE.
- Ekman, P. (1999). *Basic emotions, Handbook of cognition and emotion (pp. 45±60)*. Chichester, UK : Wiley.
- Fleiss, J. (1971). Measuring nominal scale agreement among many raters. *Psychological Bulletin*, 76(5) :378–382.
- Fourati, N. and Pelachaud, C. (2015). Relevant body cues for the classification of emotional body expression in daily actions. In *Affective Computing and Intelligent Interaction (ACII), 2015 International Conference on*, pages 267–273. IEEE.
- George, D. and Mallery, P. (2010). *SPSS for Windows Step by Step : A Simple Guide and Reference 18.0 Update*. Prentice Hall Press, Upper Saddle River, NJ, USA, 11th edition.
- Glas, N. and Pelachaud, C. (2015a). User engagement and preferences in information-giving chat with virtual agents. In *Workshop on Engagement in Social Intelligent Virtual Agents (ESIVA)*, pages 33–40.
- Glas, N. and Pelachaud, C. (2015b). User engagement and preferences in information-giving chat with virtual agents. In *Workshop on Engagement in Social Intelligent Virtual Agents (ESIVA)*, pages 33–40.

- Goutsos, D. (1997). *Modeling Discourse Topic : sequential relations and strategies in expository text*, volume 59. Greenwood Publishing Group.
- Gratch, J., DeVault, D., and Lucas, G. (2016). The benefits of virtual humans for teaching negotiation. In *International Conference on Intelligent Virtual Agents*, pages 283–294. Springer.
- Greimas, A. J. (2015). *Sémantique structurale : recherche de méthode*. Presses universitaires de France.
- Grice, H. P. (1975). Logic and conversation. 1975, pages 41–58.
- Gross, M. (1981). Les bases empiriques de la notion de prédicat sémantique. *Langages*, 63 :7–52.
- Halliday, M. (1973). *Explorations in the Functions of Language*. Explorations in Language Study. Edward Arnold.
- Hatzivassiloglou, V. and McKeown, K. R. (1997). Predicting the semantic orientation of adjectives. In *Proceedings of the 35th Annual Meeting of the Association for Computational Linguistics and Eighth Conference of the European Chapter of the Association for Computational Linguistics*, ACL '98, pages 174–181, Stroudsburg, PA, USA. Association for Computational Linguistics.
- He, R., Lee, W. S., Ng, H. T., and Dahlmeier, D. (2017). An unsupervised neural attention model for aspect extraction. In *Proceedings of the 55th Annual Meeting of the Association for Computational Linguistics (Volume 1 : Long Papers)*, volume 1, pages 388–397.
- Heider, F. (1958). *The psychology of interpersonal relations*. Lawrence Erlbaum associates Inc.
- Heylen, D., Kopp, S., Marsella, S. C., Pelachaud, C., and Vilhjálmsón, H. (2008). The next step towards a function markup language. In *International Workshop on Intelligent Virtual Agents*, pages 270–280. Springer.
- Hockett, C. F. (1958). *A course in modern linguistics*. Oxford And Ibh Publishing Co. ; New Delhi.
- Hu, M. and Liu, B. (2004). Mining opinion features in customer reviews. In *Proceedings of the 19th National Conference on Artificial Intelligence, AAAI'04*, pages 755–760. AAAI Press.
- Ishizuka, M. (2012). Textual affect sensing and affect communication. In *IEEE Transaction 11th International Conference on Cognitive Informatics and Cognitive Computing*, pages 2–3.
- Izard, C. E. (1977). *Human Emotions*. Plenum Press, New York, USA.
- Kerbrat-Orecchioni, C. (2009). *L'énonciation : de la subjectivité dans le langage*. Armand Colin.
- Klenner, M., Petrakis, S., and Fahrni, A. (2009). Robust compositional polarity classification. In *RANLP*, pages 180–184.
- Langlet, C. and Clavel, C. (2014a). Modélisation des questions de l'agent pour l'analyse des affects, jugements et appréciations de l'utilisateur dans les interactions humain-agent. In *Proceedings of TALN 2014 (Volume 2 : Short Papers)*, pages 479–484, Marseille, France. Association pour le Traitement Automatique des Langues.
- Langlet, C. and Clavel, C. (2014b). Modelling user's attitudinal reactions to the agent utterances : focus on the verbal content. In *5th International Workshop on Corpora for Research on Emotion, Sentiment & Social Signals (ES3 2014)*, Reykjavik, Iceland.
- Langlet, C. and Clavel, C. (2015a). Adapting sentiment analysis to face-to-face human-agent interactions : from the detection to the evaluation issues. In *Affective Computing and Intelligent Interaction (ACII), 2015 International Conference on*, pages 14–20. IEEE.

- Langlet, C. and Clavel, C. (2015b). Improving social relationships in face-to-face human-agent interactions : when the agent wants to know user's likes and dislikes. In *Proceedings of the 53rd Annual Meeting of the Association for Computational Linguistics and the 7th International Joint Conference on Natural Language Processing (Volume 1 : Long Papers)*, volume 1, pages 1064–1073.
- Langlet, C. and Clavel, C. (2016). Grounding the detection of the user's likes and dislikes on the topic structure of human-agent interactions. *Knowledge-Based Systems*, 106 :116–124.
- Langlet, C. and Clavel, C. (2018). Detecting user's likes and dislikes for a virtual negotiating agent. In *20th ACM International Conference on Multimodal Interaction (ICMI 2018)*, Boulder, Colorado, (Submit).
- Langlet, C., Duplessis, G. D., and Clavel, C. (2017). A web-based platform for annotating sentiment-related phenomena in human-agent conversations. In *International Conference on Intelligent Virtual Agents*, pages 239–242. Springer.
- Litman, D. J. and Forbes-Riley, K. (2004). Predicting student emotions in computer-human tutoring dialogues. In *Proceedings of the 42nd Annual Meeting on Association for Computational Linguistics*, page 351. Association for Computational Linguistics.
- Liu, B. (2012). Sentiment analysis and opinion mining. *Synthesis lectures on human language technologies*, 5(1) :1–167.
- Malandrakis, N., Potamianos, A., Iosif, E., and Narayanan, S. (2011). Emotiword : Affective lexicon creation with application to interaction and multimedia data. In *Internatinoal Workshop on computational Intelligence for Multimedia Understanding*, pages 30–41. Springer.
- Marcheggiani, D., Täckström, O., Esuli, A., and Sebastiani, F. (2014). Hierarchical multi-label conditional random fields for aspect-oriented opinion mining. In *ECIR*, pages 273–285. Springer.
- Martin, J. R. and White, P. R. (2005). *The Language of Evaluation. Appraisal in English*. Macmillan Basingstoke, London and New York.
- McKeown, G., Valstar, M., Cowie, R., Pantic, M., and Schroder, M. (2011). The semaine database : Annotated multimodal records of emotionally colored conversations between a person and a limited agent. *IEEE Transactions on Affective Computing*, 3(1) :5–17.
- Melis, L. (1988). La prédication seconde. In *Travaux de linguistique*, volume 17.
- Mel'čuk, I. and Polguère, A. (2008). Prédicats et quasi-prédicats sémantiques dans une perspective lexicographique. *Lidil. Revue de linguistique et de didactique des langues*, 37 :99–114.
- Moilanen, K. and Pulman, S. (2007). Sentiment composition. In *Proceedings of Recent Advances in Natural Language Processing (RANLP 2007)*, pages 378–382.
- Narayanan, S. and Potamianos, A. (2002). Creating conversational interfaces for children. *Speech and Audio Processing, IEEE Transactions on*, 10(2) :65–78.
- Neviarouskaya, A., Prendinger, H., and Ishizuka, M. (2010a). Emoheart : Conveying emotions in second life based on affect sensing from text. *Adv. in Hum.-Comp. Int.*, 2010.
- Neviarouskaya, A., Prendinger, H., and Ishizuka, M. (2010b). Recognition of affect, judgment, and appreciation in text. In *Proceedings of the 23rd International Conference on Computational Linguistics, COLING '10*, pages 806–814, Stroudsburg, PA, USA. Association for Computational Linguistics.

- Neviarouskaya, A., Prendinger, H., and Ishizuka, M. (2011). Sentiful : A lexicon for sentiment analysis. *IEEE Transactions on Affective Computing*, 2(1) :22–36.
- Ochs, M., Ding, Y., Fourati, N., Chollet, M., Ravenet, B., Pecune, F., Glas, N., Prepin, K., Clavel, C., and Pelachaud, C. (2013). Vers des agents conversationnels animés socio-affectifs. In *Proceedings of the 25th Conference on l'Interaction Homme-Machine*, page 69. ACM.
- Ortony, A., Clore, G., and Collins, A. (1990). *The Cognitive Structure of Emotions*. Cambridge, University Press.
- Ouali, L. O., Sabouret, N., and Rich, C. (2018). Un modèle de négociation collaborative basé sur la relation de pouvoir. In *Wacai 2018*.
- Pang, B. and Lee, L. (2005). Seeing stars : Exploiting class relationships for sentiment categorization with respect to rating scales. In *Proceedings of the 43rd annual meeting on association for computational linguistics*, pages 115–124. Association for Computational Linguistics.
- Pang, B., Lee, L., et al. (2008). Opinion mining and sentiment analysis. *Foundations and Trends® in Information Retrieval*, 2(1–2) :1–135.
- Pang, B., Lee, L., and Vaithyanathan, S. (2002). Thumbs up? : sentiment classification using machine learning techniques. In *Proceedings of the ACL-02 conference on Empirical methods in natural language processing-Volume 10*, pages 79–86. Association for Computational Linguistics.
- Paumier, S. (2011). *Unitex-manuel d'utilisation*.
- Pecune, F., Ochs, M., and Pelachaud, C. (2014). A cognitive model of social relations for artificial companions. In *International Conference on Intelligent Virtual Agents*, pages 325–328. Springer.
- Pelletier, F. J. (1994). The principle of semantic compositionality. *Topoi*, 13(1) :11–24.
- Pennebaker, J. W., Francis, M. E., and Booth, R. J. (2001). Linguistic inquiry and word count : Liwc 2001. *Mahway : Lawrence Erlbaum Associates*, 71(2001) :2001.
- Perikos, I. and Hatzilygeroudis, I. (2013). Recognizing emotion presence in natural language sentences. In *International Conference on Engineering Applications of Neural Networks*, pages 30–39. Springer.
- Poignant, J. (2016). The CAMOMILE Collaborative Annotation Platform for Multi-modal, Multilingual and Multi-media Documents. In *LREC 2016 Conference*, Portoroz, Slovenia.
- Pontiki, M., Galanis, D., Papageorgiou, H., Androutsopoulos, I., Manandhar, S., Mohammad, A.-S., Al-Ayyoub, M., Zhao, Y., Qin, B., De Clercq, O., et al. (2016). Semeval-2016 task 5 : Aspect based sentiment analysis. In *Proceedings of the 10th international workshop on semantic evaluation (SemEval-2016)*, pages 19–30.
- Qiu, G., Liu, B., Bu, J., and Chen, C. (2011). Opinion word expansion and target extraction through double propagation. *Computational linguistics*, 37(1) :9–27.
- Quirk, R. (1985). *A Comprehensive grammar of the English language*. General Grammar Series. Longman.
- Rastier, F. (1985). L'isotopie sémantique, du mot au texte. *L'Information Grammaticale*, 27(1) :33–36.

- Rentoumi, V., Petrakis, S., Klenner, M., Vouros, G. A., and Karkaletsis, V. (2010). United we stand : Improving sentiment analysis by joining machine learning and rule based methods. In *LREC*.
- Riegel, M. (2005). *Forme et interprétation des phrases copulatives à deux groupes nominaux définis : asymétrie syntaxique et configuration sémantique*. na.
- Riegel, M., Pellat, J.-C., and Rioul, R. (1994). Grammaire méthodique du français. *Linguistique nouvelle*.
- Riloff, E. and Wiebe, J. (2003). Learning extraction patterns for subjective expressions. In *Proceedings of the 2003 Conference on Empirical Methods in Natural Language Processing, EMNLP '03*, pages 105–112, Stroudsburg, PA, USA. Association for Computational Linguistics.
- Sacks, H., Schegloff, E., and Jefferson, G. (1974). A simplest systematics for the organization of turn-taking for conversation. *Language*, 50(9-10) :696–735.
- Schegloff, E. A. (1990). On the organization of sequences as a source of “coherence” in talk-in-interaction. *Conversational organization and its development*, 38 :51–77.
- Scherer, K. R. (2005). What are emotions? and how can they be measured? *Social science information*, 44(4) :695–729.
- Schuller, B., Batliner, A., Steidl, S., and Seppi, D. (2011). Recognising realistic emotions and affect in speech : State of the art and lessons learnt from the first challenge. *Speech Communication*, 53(9-10) :1062–1087.
- Searle, J. R. (1976). A classification of illocutionary acts. *Language in society*, 5(01) :1–23.
- Shaikh, M., Prendinger, H., and Ishizuka, M. (2009). A linguistic interpretation of the occ emotion model for affect sensing from text. In *Affective Information Processing*, pages 378–382. Springer London.
- Shaikh, M. A. M., Prendinger, H., and Mitsuru, I. (2007). Assessing sentiment of text by semantic dependency and contextual valence analysis. In *Affective Computing and Intelligent Interaction*, pages 191–202. Springer.
- Smith, C., Crook, N., Dobnik, S., and Charlton, D. (2011a). Interaction strategies for an affective conversational agent. In *Presence : Teleoperators and Virtual Environments*, volume 20, pages 395–411. MIT Press.
- Smith, C., Crook, N., Dobnik, S., Charlton, D., Boye, J., Pulman, S., De La Camara, R. S., Turunen, M., Benyon, D., Bradley, J., et al. (2011b). Interaction strategies for an affective conversational agent. *Presence : Teleoperators and Virtual Environments*, 20(5) :395–411.
- Socher, R., Manning, C. D., and Ng, A. Y. (2010). Learning continuous phrase representations and syntactic parsing with recursive neural networks. In *Proceedings of the NIPS-2010 Deep Learning and Unsupervised Feature Learning Workshop*, volume 2010, pages 1–9.
- Socher, R., Perelygin, A., Wu, J. Y., Chuang, J., Manning, C. D., Ng, A. Y., and Potts, C. (2013). Recursive deep models for semantic compositionality over a sentiment treebank. In *Proceedings of the conference on empirical methods in natural language processing (EMNLP)*, volume 1631, page 1642. Citeseer.
- Soleymani, M., Garcia, D., Jou, B., Schuller, B., Chang, S.-F., and Pantic, M. (2017). A survey of multimodal sentiment analysis. *Image and Vision Computing*, 65 :3–14.

- Song, Y., Pan, S., Liu, S., Zhou, M. X., and Qian, W. (2009). Topic and keyword re-ranking for lda-based topic modeling. In *Proceedings of the 18th ACM conference on Information and knowledge management*, pages 1757–1760. ACM.
- Stenetorp, P., Pyysalo, S., Topić, G., Ohta, T., Ananiadou, S., and Tsujii, J. (2012). Brat : A web-based tool for nlp-assisted text annotation. In *Proceedings of the Demonstrations at the 13th Conference of the European Chapter of the Association for Computational Linguistics, EACL '12*, pages 102–107, Stroudsburg, PA, USA. Association for Computational Linguistics.
- Su, K.-Y., Su, J., Wiebe, J., and Li, H. (2009). Proceedings of the joint conference of the 47th annual meeting of the acl and the 4th international joint conference on natural language processing of the afnlp. In *Proceedings of the Joint Conference of the 47th Annual Meeting of the ACL and the 4th International Joint Conference on Natural Language Processing of the AFNLP*.
- Szabó, Z. (2012). The case for compositionality. *The Oxford handbook of compositionality*, 64 :80.
- Taboada, M., Brooke, J., Tofiloski, M., Voll, K., and Stede, M. (2011). Lexicon-based methods for sentiment analysis. *Computational linguistics*, 37(2) :267–307.
- Turney, P. D. and Littman, M. L. (2002). Unsupervised learning of semantic orientation from a hundred-billion-word corpus. *CoRR*, cs.LG/0212012.
- Valitutti, R. (2004). Wordnet-affect : an affective extension of wordnet. In *In Proceedings of the 4th International Conference on Language Resources and Evaluation*, pages 1083–1086.
- Vion, R. (1992). La communication verbale. analyse des interactions. *Paris, Hachette*.
- Wang, W., Pan, S. J., Dahlmeier, D., and Xiao, X. (2016). Recursive neural conditional random fields for aspect-based sentiment analysis. *arXiv preprint arXiv :1603.06679*.
- Whitelaw, C., Garg, N., and Argamon, S. (2005). Using appraisal taxonomies for sentiment analysis. *Proceedings of CIKM-05, the ACM SIGIR Conference on Information and Knowledge Management*.
- Wiebe, J. (2000). Learning subjective adjectives from corpora. In *Proceedings of the Seventeenth National Conference on Artificial Intelligence and Twelfth Conference on Innovative Applications of Artificial Intelligence*, pages 735–740. AAAI Press.
- Wiebe, J. and Riloff, E. (2005). Creating subjective and objective sentence classifiers from unannotated texts. In *Proceedings of the 6th International Conference on Computational Linguistics and Intelligent Text Processing, CICLing'05*, pages 486–497, Berlin, Heidelberg. Springer-Verlag.
- Wiebe, J., Wilson, T., and Cardie, C. (2005). Annotating expressions of opinions and emotions in language. *Language resources and evaluation*, 39(2-3) :165–210.
- Wilson, T., Wiebe, J., and Hoffmann, P. (2005). Recognizing contextual polarity in phrase-level sentiment analysis. In *Proceedings of the Conference on Human Language Technology and Empirical Methods in Natural Language Processing, HLT '05*, pages 347–354, Stroudsburg, PA, USA. Association for Computational Linguistics.
- Yang, B. and Cardie, C. (2013). Joint inference for fine-grained opinion extraction. In *Proceedings of the 51st Annual Meeting of the Association for Computational Linguistics (Volume 1 : Long Papers)*, pages 1640–1649, Sofia, Bulgaria. Association for Computational Linguistics.
- Yildirim, S., Narayanan, S., and Potamianos, A. (2011). Detecting emotional state of a child in a conversational computer game. *Computer Speech & Language*, 25(1) :29–44.

- Zhang, L. (2012a). *Automatic opinion analysis : issue of intensity and negation in a newspaper corpus*. Theses, Université de Caen.
- Zhang, L. (2012b). *Automatic opinion analysis : issue of intensity and negation in a newspaper corpus*. PhD thesis, Université de Caen.
- Zhang, T., Hasegawa-Johnson, M., and Levinson, S. E. (2006). Cognitive state classification in a spoken tutorial dialogue system. *Speech communication*, 48(6) :616–632.
- Zhuang, L., Jing, F., and Zhu, X.-Y. (2006). Movie review mining and summarization. In *Proceedings of the 15th ACM international conference on Information and knowledge management*, pages 43–50. ACM.

# Appendices



## Annexe A

# Glossaire

- Affect : sous-classe d'attitude qui regroupe les émotions et les réactions affectives.
- Appréciation : sous-classe d'attitude qui regroupe les évaluations d'artefacts, d'événements ou de processus.
- Attitude : classe référant, selon la terminologie de Martin and White (2005), aux valeurs que les locuteurs communiquent et aux réponses émotionnelles ou affectives qu'ils associent aux participants et aux processus. Le modèle de Martin and White (2005) compte trois classes d'attitude : les affects, les appréciations et les jugements.
- Attribut du sujet : deuxième constituant d'un verbe d'état permettant d'exprimer une caractérisation (propriété, état ou catégorisation) (Riegel, 2005).
- Attribut de l'objet : complément de l'objet qui entretient avec celui-ci le même rapport que l'attribut du sujet avec le sujet (Riegel et al., 1994).
- Axiologique : qui porte un jugement de valeur. Dans le langage, une catégorisation axiologique est portée par un mot cumulant une description du dénoté et une évaluation qui se réfère à une norme sociale et morale (Kerbrat-Orecchioni, 2009).
- Compositionnalité sémantique : principe selon lequel l'interprétation d'une expression complexe est une fonction de l'interprétation de ses parties et de la manière dont elles sont assemblées (Pelletier, 1994).
- Isotopie sémantique : récurrence d'un même trait sémantique au sein d'un énoncé, d'un texte ou d'un discours. L'isotopie sémantique participe ainsi de l'homogénéité du discours (Rastier, 1985), (Greimas, 2015).
- Jugement : sous-classe d'attitude qui regroupe des évaluations de comportements ou d'êtres animés.
- *Like* ou *Dislike* : sous-classe transverse d'attitude regroupant toutes les expressions d'attitude informant sur les goûts de l'utilisateur (voir Chapitre 6).
- Expression simple d'attitude : dans une conversation, expression d'attitude dont tout ou un partie du sens peut être appréhendé par la connaissance seule de l'un des tours de parole composant une paire adjacente.
- Expression d'attitude strictement collaborative : dans une conversation, expression d'attitude dont le sens ne peut être appréhendé que par une connaissance des deux tours de parole composant une paire adjacente.
- Paire adjacente : ensemble de deux tours de parole produits par deux locuteurs différents reliés par une dépendance conditionnelle. Cette notion permet de décrire l'organisation séquentielle de l'échange dans une conversation (Sacks et al., 1974).

- Prédicat sémantique : unité lexicale dénotant un fait pouvant être un évènement, une action, une activité, un état, un processus, une relation, etc. Le prédicat a un sens liant qui présuppose nécessairement un certain nombre de participants (Mel'čuk and Polguère, 2008), (Gross, 1981).
- Procès : sens dénoté par un verbe, que celui-ci exprime un fait, un événement, un état ou tout autre notion (Riegel et al., 1994).

## Annexe B

# Plateforme pour l'annotation du corpus Semaine

Dans cette annexe, nous présentons la plateforme développée pour la campagne d'annotation du corpus Semaine. Toutes les informations concernant le protocole d'annotation peuvent être trouvées Chapitre 7. Les figures ci-dessous présentent les questionnaires intégrés à la plateforme. Chacun d'eux se concentre sur l'une des propriétés de l'expression de *like* ou *dislike* détectée : polarité, identité de la cible, mots référant à la cible dans le tour de parole de l'agent et dans celui de l'utilisateur et enfin intensité. Le jeu d'annotation obtenu via cette plateforme dans le cadre de la campagne d'annotation Semaine est décrit en Annexe C.

---

---

**A:** *"What do you think about them?"*

**B:** *"They are wonderful!"*

---

---

**QUESTION 1: YOUR TURN!**

According to these sentences, do you think that the B speaker likes/dislikes something?

yes no

FIGURE B.1 – Questionnaire 1 : la présence d'un *like* ou *dislike* dans la paire considérée

---

---

**A:** "What do you think about them?"

**B:** "They are wonderful!"

---

---

QUESTION 2: YOUR TURN!

Does the B speaker's sentence contain more than one expression of what he likes/dislikes

yes  no

FIGURE B.2 – Questionnaire 2 : nombre d'expressions de *like* ou *dislike* présentes dans la paire considérée

---

---

**A:** "What do you think about them?"

**B:** "They are wonderful!"

---

---

QUESTION 3: YOUR TURN!

Does the B speaker like/dislike

the A speaker  the A speaker's sentence  the B speaker  something else  we cannot guess the nature of the thing that he likes/dislikes

FIGURE B.3 – Questionnaire 3 : polarité

---

QUESTION 4: YOUR TURN!

In these sentences, select the word(s) corresponding to the thing (or the set of things) that the B speaker likes/dislikes

THE A SPEAKER'S SENTENCE

What	do	you	think	about	them	No word corresponding to the thing (or set of things) liked/disliked in this sentence
<input type="checkbox"/>						

THE B SPEAKER'S SENTENCE

They	are	wonderful	No word corresponding to the thing (or set of things) liked/disliked in this sentence
<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>

FIGURE B.4 – Questionnaire 4 : type de la cible

---

A: "What do you think about them?"

B: "They are wonderful!"

---

QUESTION 5: YOUR TURN!

Does the user :

<input type="radio"/>	<input type="radio"/>
like this	dislike this

FIGURE B.5 – Questionnaire 5 : sélection des mots, dans le tour de parole de l'agent et le tour de parole de l'utilisateur, référant à la cible

---

---

**A:** *"What do you think about them?"*

**B:** *"They are wonderful!"*

---

---

QUESTION 6: YOUR TURN

How strong is your belief about all the answers you have chosen for the sentences ?

very weak                        very strong

FIGURE B.6 – Questionnaire 6 : intensité

## Annexe C

# Jeu d'annotation : *likes* et *dislikes* dans le corpus Semaine

Au terme de la campagne d'annotation menée via *Amazon Mechanical Turk* et décrite dans le chapitre 7 7.2, nous avons disposé d'un jeu d'annotation concernant une partie du corpus Semaine. Seules 600 paires ont été annotées. Les informations relatives au jeu d'annotations sont réparties dans trois tables. La table C.1 donne pour chaque paire, son identifiant unique la session d'origine, le *time code*<sup>1</sup> de début du premier mot, le *time code* de fin du dernier mot, et l'identifiant de chacun des 4 annotateurs l'ayant annotée. La table C.2 donne pour chaque paire le nombre d'expressions de *like* ou *dislike* que chaque annotateur a trouvé. La table C.3 donne pour chaque expression trouvée, dans chaque paire et pour chaque annotateur, la polarité, le type de la cible et les mots référant à la cible dans le tour de parole de l'agent et celui de l'utilisateur, et enfin l'intensité.

---

1. Les *time codes* donnés dans la première table correspondent à ceux que l'on peut trouver dans le corpus Semaine et peuvent être trouvés via le lien [semaine-db.eu](http://semaine-db.eu)

Pair ID	Session	Start	End	Annotator 1	Annotator 2	Annotator 3	Annotator 4
1	session079	374330	379170	54acf9607b437	54acf93d0882c	54acfb326c9c9	54ad0b80ce452
2	session079	380110	382340	54acf9607b437	54acf93d0882c	54acfb326c9c9	54ad0b80ce452
3	session079	390050	392070	54acf9607b437	54acf93d0882c	54acfb326c9c9	54ad0b80ce452
4	session079	452250	453020	54acf93d0882c	54acf9607b437	54acfb326c9c9	54ad0b80ce452
5	session082	162600	163200	54acf9607b437	54acf93d0882c	54acfb326c9c9	54ad0b80ce452
6	session082	169420	170010	54acf9607b437	54acf93d0882c	54acfb326c9c9	54ad0b80ce452
7	session082	171900	172740	54acf9607b437	54acf93d0882c	54acfb326c9c9	54ad0b80ce452
8	session082	173150	174420	54acf9607b437	54acf93d0882c	54acfb326c9c9	54ad0b80ce452
9	session082	179710	180990	54acf9607b437	54acf93d0882c	54acfb326c9c9	54ad0b80ce452
10	session082	183820	185470	54acf9607b437	54acf93d0882c	54acfb326c9c9	54ad0b80ce452
11	session082	187070	190850	54acf97f03a3f	54acfa5fac166	54acfa0467bb4	54acfe23c962a
12	session085	3600	4650	54acf97f03a3f	54acfa5fac166	54acfa0467bb4	54acfe23c962a
13	session085	5320	5950	54acf97f03a3f	54acfa5fac166	54acfa0467bb4	54acfe23c962a
14	session085	7550	8240	54acf97f03a3f	54acfa5fac166	54acfa0467bb4	54acfe23c962a
15	session085	9950	10770	54acf97f03a3f	54acfa5fac166	54acfa0467bb4	54acfe23c962a
16	session085	12210	13180	54acf97f03a3f	54acfa5fac166	54acfa0467bb4	54acfe23c962a
17	session085	14320	15310	54acf97f03a3f	54acfa5fac166	54acfa0467bb4	54acfe23c962a
18	session085	37270	38230	54acf97f03a3f	54acfa5fac166	54acfa0467bb4	54acfe23c962a
19	session085	38870	39250	54acf97f03a3f	54acfa5fac166	54acfa0467bb4	54acfe23c962a
20	session085	76740	80530	54acf97f03a3f	54acfa5fac166	54acfa0467bb4	54acfe23c962a
21	session085	84150	84250	54afd33bdfdbd	54afd7dd412b0	54afd9f21efd5	54b3fc5abddf0
22	session085	85780	86480	54afd33bdfdbd	54afd7dd412b0	54afd9f21efd5	54b3fc5abddf0
23	session085	88110	88850	54afd33bdfdbd	54afd7dd412b0	54afd9f21efd5	54b3fc5abddf0
24	session085	90910	91350	54afd33bdfdbd	54afd7dd412b0	54afd9f21efd5	54b3fc5abddf0
25	session085	107400	110340	54afd33bdfdbd	54afd7dd412b0	54afd9f21efd5	54b3fc5abddf0
26	session085	112050	113050	54afd33bdfdbd	54afd7dd412b0	54afd9f21efd5	54b3fc5abddf0
27	session085	117010	117650	54afd33bdfdbd	54afd7dd412b0	54afd9f21efd5	54b3fc5abddf0

FIGURE C.1 – Aperçu de la table contenant les informations relatives aux paires (identifiant de la paire, session d’origine et annotateurs)

Pair ID	Annotation round ID	Annotator	Nb Attitude Per Pair
1	2	54acf9607b437	1
1	23	54acf93d0882c	1
1	58	54acfb326c9c9	1
1	80	54ad0b80ce452	1
2	8	54acf9607b437	0
2	26	54acf93d0882c	1
2	53	54acfb326c9c9	1
2	81	54ad0b80ce452	1
3	6	54acf9607b437	0
3	14	54acf93d0882c	1
3	62	54acfb326c9c9	1
3	76	54ad0b80ce452	0
4	3	54acf93d0882c	1
4	5	54acf9607b437	0
4	61	54acfb326c9c9	0
4	75	54ad0b80ce452	0
5	10	54acf9607b437	0
5	35	54acf93d0882c	1
5	54	54acfb326c9c9	0
5	79	54ad0b80ce452	0
6	7	54acf9607b437	0
6	47	54acf93d0882c	0
6	60	54acfb326c9c9	1
6	74	54ad0b80ce452	1
7	1	54acf9607b437	1
7	45	54acf93d0882c	0
7	57	54acfb326c9c9	1

FIGURE C.2 – Aperçu de la table indiquant le nombre d'expressions de *like* ou *dislike* trouvées dans chaque paire par chaque annotateur

ID	Annotation Round	Target Type	Target Words Agent	Target Word User	Intensité	
23	7	positive	other	_time/	_wait/_yeah/_it/_exciting	4
26	8	positive	other	_noTargetWd	_l/_aunty/_nice	5
28	9	positive	agent	_noTargetWd	_you/	3
29	10	negative	other	_goodness/_win/	_don/_t/	5
41	12	positive	other	_congratulations/	_thank/_thank/	4
44	13	positive	agent	_noTargetWd	_Prudence/	4
53	14	positive	other	_noTargetWd	_aunty/_aunty_it	3
57	15	positive	unknown	_that/	_noTargetWd	2
58	16	positive	unknown	_noTargetWd	_it/_it	3
59	17	positive	unknown	_noTargetWd	_noTargetWd	3
60	18	positive	unknown	_that/	_noTargetWd	3
62	19	positive	other	_noTargetWd	_one/	2
73	20	positive	unknown	_that/	_noTargetWd	3
74	21	positive	unknown	_that/	_noTargetWd	3
80	22	positive	other	_noTargetWd	_it	4
81	23	positive	other	_noTargetWd	_nice	4
100	32	positive	other	_noTargetWd	_noTargetWd	2
103	31	negative	unknown	_noTargetWd	_feel/_upset/_silly/	4
118	34	positive	unknown	_that/	_noTargetWd	3
126	36	positive	unknown	_noTargetWd	_things/	1
127	35	negative	unknown	_angry/	_yeah/	2
131	36	positive	other	_noTargetWd	_bi-lingual/_studies/	4
132	37	positive	agent	_noTargetWd	_Prudence/	5
133	38	positive	other	_sunshine/	_precisely/	3
136	39	negative	other	_impulse/_shopping/	_precisely/	3

FIGURE C.3 – Caractéristiques de chaque expression de *like* et *dislike* identifiée par chaque annotateur

## Annexe D

# A Web-Based Platform for Annotating Sentiment Related Phenomena in Human-Agent Conversations

Article de démonstration présenté à IVA 2017. Langlet, C., Duplessis, G. D., and Clavel, C. (2017). A web-based platform for annotating sentiment-related phenomena in human-agent conversations. In *International Conference on Intelligent Virtual Agents*, pages 239–242. Springer

### D.1 Introduction

In the research field of embodied conversational agents, the detection of sentiment-related phenomena is a challenging task contributing to the improvement of human-agent interactions. Verbal content is more and more integrated (Langlet and Clavel, 2016). Developing a detection system of verbal expressions of sentiment in human-agent conversations requires annotated corpora. Such corpora are rare and managing annotation campaigns appears as the first step in the development of a detection system. In order to make easier the implementation of datasets, we have created a web-based platform focusing on the verbal expression of sentiment and suited to transcriptions of human-agent conversations.

### D.2 A Platform Focusing on Verbal Content and Conversational Structure

**Focus on Verbal Content.** Our annotation platform deals with verbal content, without non-verbal context. The annotator has only access to manually transcribed conversations. In this way, we aim to get information about the verbal form of the sentiment-related phenomena, for helping the design and evaluation of algorithms grounded on either machine learning or rule-based models.

**Dealing with Conversational Structure.** To the best of our knowledge, the sentiment analysis research does not provide annotation sets in human-agent conversation context. The existing annotation tools are mostly designed for non-conversational texts and do not deal with the specific features of conversation (see (Poignant, 2016) and (Stenetorp et al., 2012)). In order to fill this gap, we give attention to two important dialogue features : adjacency pairs and conversation progression. First, adjacency pairs are defined as the minimal annotation context (see Area 2 in Figure D.1). Second, the annotator has the previous speech turns and their annotation (see Area

1 in Figure D.1). She/he can keep in mind past information of the conversation progress while annotating a specific adjacency pair.

**Simplifying the Annotation Task.** Annotating sentiment-related phenomena is a challenging task. It depends on the personal subjectivity and interpretation of verbal content. This platform provides a frame which guides the annotator and simplifies the annotation process. No span of text or complex semantic features have to be defined : each adjacency pair is displayed with questionnaires asking information about sentiment expressions.

**A Generic Framework.** The framework relies on standard web technologies (PHP, HTML, JAVASCRIPT) and can deal with various dyadic human-agent conversation corpora. For using our annotation platform, a simple pre-processing of the data is needed to : (i) uniquely identify conversations (via a unique ID), (ii) turn each conversation into a sequence of adjacency pairs (first, the agent’s speech turn, then the user’s speech turn), (iii) in the transcription, no paralinguistic annotations are allowed, (iv) a specific set of targets must be defined (the targets can be either artefacts or animated entities).

### D.3 Annotating User’s Sentiments and Preferences

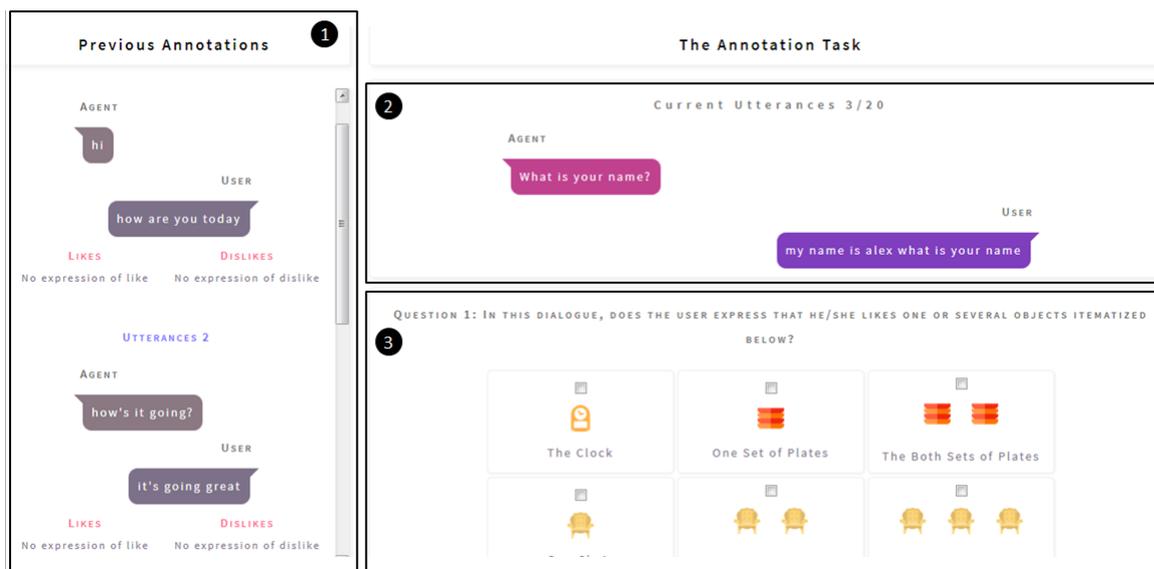


FIGURE D.1 – Platform Overview

The platform deals with two different annotation tasks which are both key tasks in the research field of virtual agent (see (Clavel and Callejas, 2016) and (Glas and Pelachaud, 2015b)) : (i) the detection of sentiments related-phenomena toward specific objects; (ii) the ranking of user’s preferences. As we consider that it is too difficult for a single annotator to concurrently decide which object is the target of a sentiment and which one is the most liked, the two tasks are strictly independent from each other and are assigned to different annotators. For each task<sup>1</sup>, the framework is divided in two parts : first, the current annotation task which concerns a specific pair (Area 2 and 3 in Fig. D.1) ; second, the history (Area 1 in Fig. D.1) which comprises all the previous pairs and the related annotations.

1. perso.telecom-paristech.fr/ langlet/annotationNegociation/task1, the task 2 at perso.telecom-paristech.fr/ langlet/annotationNegociation/task2

**Task 1.** At each adjacency pair, the annotator has to detect an expression of sentiment targeting specific objects by filling in a questionnaire, split in two questions. The first question (see Figure D.3) concerns presence of positive sentiment-related phenomena, the second, the presence of negative sentiment-related phenomena.

QUESTION 1: IN THIS DIALOGUE, DOES THE USER EXPRESS THAT HE/SHE LIKES ONE OR SEVERAL OBJECTS INVOLVED IN THE NEGOCIATION AND ITEMATIZED BELOW?

FIGURE D.2 – First questionnaire of Task 1

The user’s sentiments related phenomena toward an object may change during the conversation. As the task is to find expressions of sentiment within a specific pair, the annotator can check a positive expression of sentiment toward an object even if a negative sentiment toward the same object was previously noted. In this way, we will be able to analyse user’s sentiments along the conversation.

**Task 2.** For ranking preferences, each object needs to be scored. The score depends on the number of objects comprised in the set. In the demonstration, we use a score from 1 to 3 because the annotator has to rank preferences toward three different objects. Each score indicates the position of the object in the scale of the user’s preferences. For example, when the user verbally expresses a preference for a painting over records and lamps, the annotator can attribute to the painting the score 3 and to the lamps and the records, the score 1. Similarly to Task 1, user’s preferences may change during the conversation. As an annotation is asked for each pair, the annotator will be able to make modifications, that could be appropriate, to any score. In this way, we can follow the evolution of user’s preferences throughout the conversation.

FIGURE D.3 – First questionnaire of Task 1



## Annexe E

# Implémentation des grammaires dans *Unitex*

### E.1 Les grammaires hors-contexte

Les grammaires que nous utilisons dans notre système sont des variantes des grammaires hors-contexte. Une grammaire est composée de règles de réécriture. Celles-ci formalisent les différentes séquences que la grammaire peut prendre en compte. Le formalisme traditionnel place la règle à la suite d'une flèche  $\rightarrow$ . Par exemple, la grammaire ci-dessous est composée de deux règles de réécriture.

$$\begin{aligned} V &\rightarrow \text{"dislike"} \\ V &\rightarrow \text{"like"} \end{aligned}$$

Les symboles terminaux sont des littéraux – des mots par exemple – qui ne peuvent être subdivisés en éléments plus petits. Ils peuvent pas être changés par les règles de réécriture et ne peuvent apparaître qu'à droite de la flèche (dans l'exemple ci-dessous, "like" et "dislike"). Les symboles non-terminaux – également appelés variables syntaxiques – peuvent être remplacés et changés par des règles de réécriture. Dans l'exemple ci-dessus, ils apparaissent à gauche de la flèche ( $V$ ). Dans nos grammaires, les symboles non-terminaux sont les labels attribués aux formes syntaxiques repérées ( $SAdj$ ,  $SV$ ,  $SN$ ).

### E.2 Grammaires et graphes dans *Unitex*

Les grammaires *Unitex*<sup>1</sup> sont appliquées comme des transducteurs, c'est-à-dire qu'elles peuvent produire des sorties (ajouter des labels au texte, par exemple).

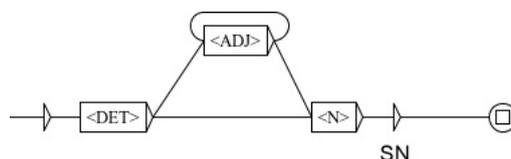


FIGURE E.1 – Exemple de graphe Unitex

Afin de faciliter leur écriture, *Unitex* propose une interface représentant les grammaires sous forme de graphes visuels (Figure E.1). La flèche représente l'état initial, le rond contenant un carré représente l'état final. Entre l'état initial et l'état final, se trouve un ensemble de chemins

1. Pour plus d'informations sur la plateforme *Unitex*, voir le manuel de l'utilisateur (Paumier, 2011)

jalonnés par des boîtes. Ces chemins représentent les séquences pouvant être reconnues par la grammaire. Une boîte représente un symbole terminal ou non-terminal composant la séquence. Aux boîtes peuvent être associées des sorties. Nous avons utilisé ces sorties pour associer des labels aux séquences reconnues par la grammaire. Un graphe peut faire appel soit à un sous graphe, soit aux annotations d'un graphe appliqué antérieurement dans le processus de traitement. Un graphe peut contenir des boucles. Le contenu de chaque boucle est reconnu n'importe quel nombre de fois. Le graphe représenté Figure E.1 compte deux chemins. Le premier permet de reconnaître une séquence faisant suivre un déterminant (*<DET>*) par un nom (*<N>*), le second une séquence se composant d'un déterminant, d'un adjectif et d'un nom. La boucle indique que la grammaire peut reconnaître autant d'adjectifs que nécessaire. Les non-terminaux *<ADJ>*, *<N>* et *<DET>* sont les labels produits par les traitements antérieurs. Le graphe produit une sortie, le label *<SN>* : il sera associé à toutes les séquences reconnues par le graphes.

La figure E.2 montre un autre exemple de graphe plus complexe et permettant de repérer quatre variantes syntaxiques d'un énoncé de type : *<Source> <Feel> <Affect> <Target>*. Les différents chemins nous permettent d'envisager différentes configurations syntaxiques : verbe porteur d'une négation (*<FEEL+mod :neg>*) suivi d'un adjectif ayant une polarité soit positive, soit négative. Pour chacune de ces configurations, nous sommes ainsi capable de définir la polarité de sortie de l'expression repérée. Le label *L&D* placé en sortie de graphe est attribué à l'expression repérée ainsi que la structure de trait associée ( $pol = \$Pol\$ + target = \$Target\$ + source = User$ ). L'attribut *pol* représente la polarité de l'expression, sa valeur est attribuée sur la base des informations morpho-syntaxiques et sémantiques des éléments repérés par le graphe. Le groupe placé entre les (*TARGET*)*TARGET* est considéré comme référant à la cible de l'expression et son littéral (ou chaîne de caractère) est récupéré et assigné comme valeur à l'attribut *target*.

### E.3 Cascade de transducteurs

Les grammaires développées ont été appliquées en cascade. Une cascade de grammaires applique successivement plusieurs graphes sur le texte. Chaque grammaire apporte des transformations sur le texte : ajout de labels ou regroupement de plusieurs unités qui constitueront une seule et même unité pour les traitements postérieurs. Tous les changements effectués par un graphe peuvent être employés par les graphes suivants. Dans notre système, une cascade de grammaires permet de mettre en place l'analyse *bottom-up* présentée Chapitre 6 (analyse lexicale, analyse syntaxique et analyse phrastique). Le tableau E.1 présente la cascade de grammaires utilisées dans le cadre d'interaction *small-talk*. Pour chacune des grammaires, il détaille les sorties qu'elle génère. Le détail des structures syntaxiques qu'elles peuvent repérer est présenté Annexe F.

L'interface graphique proposée par *Unitex* n'a été utilisée que pour le développement des grammaires. Pour leur application sur le texte, nous avons utilisé *python-unitex*. Ce package donne accès à tous les fonctionnalités de la librairie C++ d'*Unitex* via un programme python.

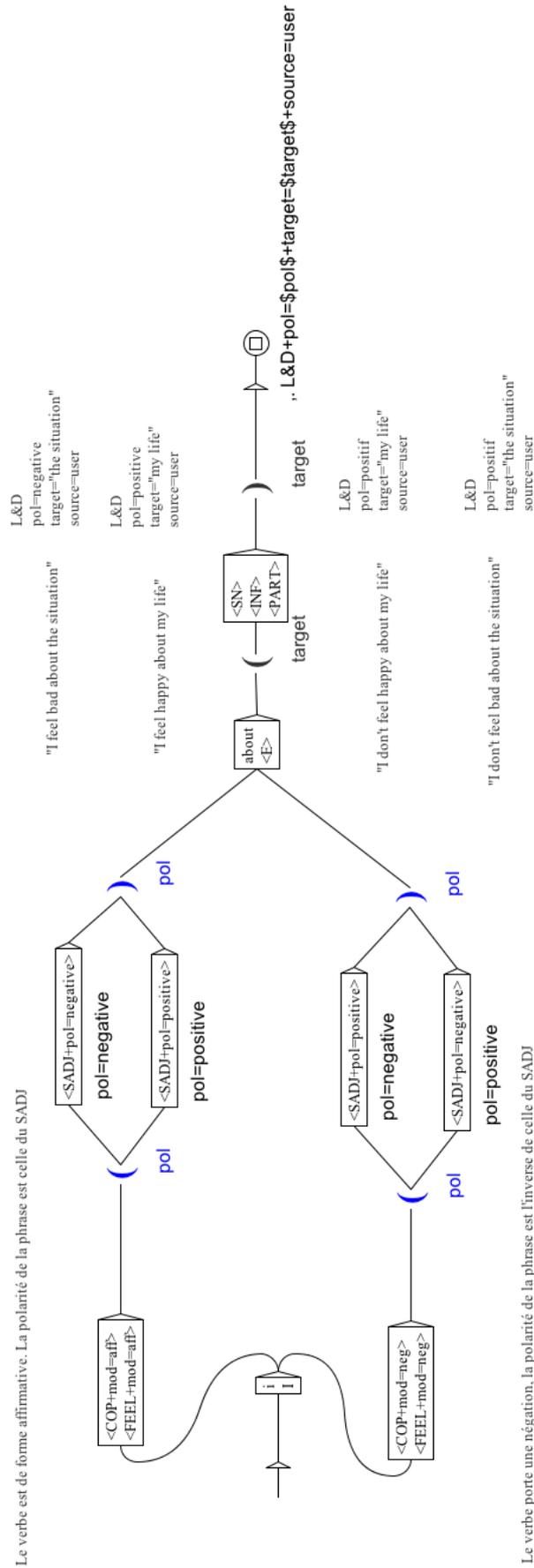


FIGURE E.2 – Graphe *Unitex* permettant de détecter des expressions de *like* ou *dislike* de type "I feel happy about life"

Grammaires de niveau lexical	
auxiliaires.grf	Repère des auxiliaires verbaux
opinionVb.grf	Repère des verbes synonymes de <i>see</i> et pouvant être utilisés dans les structures de type <i>I see as beautiful. Consider, see, find, regard</i>
attributifVb.grf	Repère des verbes pouvant jouer le rôle de verbe attributif dans une structure de type <i>The painting is beautiful</i>
Grammaires du niveau syntagmatique	
SAdj.grf	Repère les syntagmes adjectivaux
SN.grf	Repère les syntagmes nominaux
SV-Attributif.grf	Traite les syntagmes verbaux dont le verbe principal est un verbe attributif ayant été repéré par la grammaire <i>verbsAttributifsListe.grf</i>
SV-opinion.grf	Traite les syntagmes verbaux dont le verbe principal est un verbe d'opinion repéré par <i>opinionVb.grf</i>
SV-L&D	Repère les syntagmes verbaux exprimant un <i>like</i> ou un <i>dislike</i> et dont le verbe principal est issu de <i>WordNetAffect</i> )
SV.grf	Repère les syntagmes verbaux non pris en charge par les précédentes grammaires
SN-etendu.grf	Repère des syntagmes nominaux auxquels est rattachée une proposition relative
InfParticipiale.grf	Repère des propositions infinitives ou participiales
Grammaires de niveau phrastique	
L&D-Feel	Reconnait les structures de type <i>I feel terrible about that</i>
L&D-Make	Reconnait les structures de type <i>This movie makes me sad</i>
L&D-Have	Reconnait les structures de type <i>I had an awful week</i>
L&D-Attributif	Reconnait les structures de type <i>Her brother was so brave</i>
L&D-Pred	Reconnait les structures de type <i>I don't like this weather</i>
L&D-Pred	Reconnait les structure de type <i>I don't like this weather</i>

 TABLEAU E.1 – Cascade de grammaires *Unitex* utilisée par notre système

## Annexe F

# Détail des grammaires

Les tableaux ci-dessous présentent en détail les principales grammaires utilisées dans notre chaîne de traitement pour le *small-talk* et listées Annexe D. Les chemins qui composent chaque grammaire sont représentés avec le formalisme plus traditionnel utilisé pour la grammaire hors-contexte. Chaque séquence suivant le symbole  $\rightarrow$  représente un chemin, c'est-à-dire une structure syntaxique pouvant être repérée par la grammaire. L'ensemble des symboles utilisés dans les grammaires est décrit dans le tableau F.1 ci-dessous. Ces symboles peuvent être accompagnés de structure de traits mises entre parenthèses.

Symbole	Définition
Det	Déterminant
N(pol :[neutral positive negative])	Nom de polarité positive, négative ou neutre
SN(pol :[neutral positive negative])	Syntagme nominal de polarité positive, négative ou neutre
Sadj(pol :[neutral positive negative])	Syntagme adjectival de polarité positive, négative ou neutre
SV(type :L&D, pol :[positive negative])	Syntagme verbal exprimant un <i>like</i> ou un <i>dislike</i> de polarité positive ou négative
Adv	Adverbe
INF	Proposition Infinitive
PART	Proposition Participiale
lemme :L	dont le lemme est L
negWord(type :pre)	mot de négation préposé au verbe
negWord(type :post)	mot de négation postposé au verbe
be(mod :aff)	verbe ayant une modalité affirmative. Ex : <i>Like</i>
be(mod :neg)	verbe ayant une modalité négative. Ex : <i>don't like</i>
tag :vb	unité lexicale ayant été étiquetée comme verbe par le POS-Tagger d'unitex
vb("ing")	verbe se terminant par "ing"

TABLEAU F.1 – Liste des symboles utilisés dans les grammaires

### F.1 Grammaires de niveau syntagmatique

<p>SV(type :opinion, mod :neg)</p> <p>↑</p>	<p>auxiliaire(mod :neg), Adv ?, SV(type :opinion, mod :aff, voice :active)</p> <p>will(mod :neg), Adv ?, negWord(type :post), vb(lemme :have), "been", vbOpinion("ing")</p> <p>will(mod :neg), Adv ?, negWord(type :post), vb(lemme :have), vbOpinion(partPast)</p> <p>will(mod :neg), Adv ?, negWord(type :post), be(mod :aff), Adv ?, vbOpinion("ing")</p> <p>will(mod :neg), Adv ?, negWord(type :post), vbOpinion</p> <p>vb(lemme :have), Adv ?, negWord(type :post), "been", vbOpinion("ing")</p> <p>vb(lemme :have), Adv ?, negWord(type :post), vbOpinion(partPast)</p> <p>be(mod :neg), Adv ?, vbOpinion("ing")</p> <p>negWord(type :pre), vbOpinion</p>	<p>"shouldn't consider"</p> <p>"will not have been considering"</p> <p>"will not have considered"</p> <p>"will not be considering"</p> <p>"will not consider"</p> <p>"have not been considering"</p> <p>"haven't considered"</p> <p>"am not only considering"</p> <p>"don't consider"</p>
<p>SV(type :opinion, mod :aff, voice :active)</p> <p>↑</p>	<p>will, Adv ?, vb(lemme :have), "been", vbOpinion("ing")</p> <p>will, Adv ?, vb(lemme :have), vbOpinion(partPast)</p> <p>will, be(mod :aff), Adv ?, vbOpinion("ing")</p> <p>will, be(mod :aff), Adv ?, vbOpinion</p> <p>vb(lemme :have), "been", vbOpinion("ing")</p> <p>vb(lemme :have), vbOpinion(partPast)</p> <p>be(mod :aff), Adv ?, vbOpinion("ing")</p> <p>vbOpinion</p>	<p>"should consider"</p> <p>"will have considered"</p> <p>"will be considering"</p> <p>"will consider"</p> <p>"have been considering"</p> <p>"have considered"</p> <p>"am only considering"</p> <p>"consider"</p>
<p>SV-opinion.grf</p>	<p>vbOpinion</p> <p>vbOpinion</p> <p>vbOpinion</p> <p>vbOpinion</p>	<p>vb(lemme :see)</p> <p>vb(lemme :consider)</p> <p>vb(lemme :find)</p> <p>vb(lemme :regard)</p>

TABLEAU F.2 – Grammaires traitant les syntagmes verbaux intégrant un verbe d’opinion (to think, to feel, to see, etc)

SV-L&D.grf	SV(type :L&D, mod :neg, pol :inverse( <i>PolVbLikeDislike</i> ))	→	auxiliaire(mod :neg), Adv?, SV(type :L&D, mod :aff)	“shouldn't love”
		→	will(mod :neg), Adv?, negWord(type :post), vb(lemme :have), "been", vbLikeDislike("ing")	“will not have been loving”
		→	will(mod :neg), Adv?, negWord(type :post), vb(lemme :have), "had"	“will not have been loving”
		→	will(mod :neg), Adv?, negWord(type :post), be(mod :aff), Adv?, vbLikeDislike("ing")	“will not be loving”
		→	vb(lemme :have), Adv?, negWord(type :post), "been", vbLikeDislike("ing")	“have not been loving”
		→	vb(lemme :have), Adv?, negWord(type :post), vbLikeDislike(partPast)	“haven't loved”
		→	be(mod :neg), Adv?, vbLikeDislike("ing")	“am not only loving”
		→	negWord(type :pre), vbLikeDislike	“don't love”
	SV(type :L&D, mod :aff, pol : <i>PolVbLikeDislike</i> )	→	auxiliaire(mod :pos), Adv?, SV(type :L&D, mod :aff, voice :active)	“could love”
		→	will, Adv?, vb(lemme :have), "been", vbLikeDislike("ing")	will have been loving
		→	will, Adv?, vb(lemme :have), vbLikeDislike(partPast)	“will have loved”
		→	will, be(mod :aff), Adv?, vbLikeDislike("ing")	“will be loving”
		→	vb(lemme :have), "been", vbLikeDislike("ing")	“have been loving”
		→	vb(lemme :have), vbLikeDislike(partPast)	“have loved”
	→	be(mod :aff), Adv?, vbLikeDislike("ing")	“am loving”	
	→	vbLikeDislike	“love”	
	→	vb(inWordNet-Affect)		

TABLEAU F.3 – Grammaire traitant les syntagmes verbaux exprimant un *like* ou un *dislike*

SV-Attributif.grf	SV(type :attributif, mod :neg)	<p>↑ auxiliare(mod :neg), Adv?, SV(type :attributif, mod :aff)</p> <p>↑ will(mod :neg), Adv?, negWord(type :post), vb(lemme :have), "been", vbAttributif("ing")</p> <p>↑ will(mod :neg), Adv?, negWord(type :post), vb(lemme :have), vbAttributif("ing")</p> <p>↑ will(mod :neg), Adv?, negWord(type :post), be(mod :aff), Adv?, vbAttributif("ing")</p> <p>↑ vb(lemme :have), Adv?, negWord(type :post), "been", vbAttributif("ing")</p> <p>↑ vb(lemme :have), Adv?, negWord(type :post), vbAttributif(partPast)</p> <p>↑ be(mod :neg), Adv?, vbAttributif("ing")</p> <p>↑ negWord(type :pre), vbAttributif</p>	<p>"must not be"</p> <p>"will not have been being"</p> <p>"will not have been being"</p> <p>"will not be being"</p> <p>"have not been being"</p> <p>"haven't been"</p> <p>"isn't being"</p> <p>"am not"</p>
SV-Attributif.grf	SV(type :attributif, mod :aff)	<p>↑ auxiliare(mod :pos), Adv?, SV(type :attributif, mod :aff)</p> <p>↑ will, Adv?, vb(lemme :have), "been", vbAttributif("ing")</p> <p>↑ will, Adv?, vb(lemme :have), vbAttributif(partPast)</p> <p>↑ will, be(mod :aff), Adv?, vbAttributif("ing")</p> <p>↑ vb(lemme :have), "been", vbAttributif("ing")</p> <p>↑ vb(lemme :have), vbAttributif(partPast)</p> <p>↑ be(mod :aff), Adv?, vbAttributif("ing")</p> <p>↑ vbAttributif</p>	<p>"could be"</p> <p>will have been being</p> <p>"will have been"</p> <p>"will be being"</p> <p>"have been being"</p> <p>"have been"</p> <p>"is being"</p> <p>"am"</p>
attributifVb.grf	vbAttributif	<p>↑ vb(lemme :be)</p> <p>↑ vb(lemme :seem)</p> <p>↑ vb(lemme :get)</p> <p>↑ vb(lemme :appear)</p> <p>↑ vb(lemme :sound)</p> <p>↑ vb(lemme :look)</p>	

TABLEAU F.4 – Grammaires traitant les syntagmes verbaux attributifs

<p>SV:grf</p>	<p>SV(type :vb, mod :neg)</p>	<p>→ → → → → → → →</p>	<p>auxiliaire(mod :neg), Adv?, SV(type :vb, mod :aff) will(mod :neg), Adv?, negWord(type :post), vb(lemme :have), "been", vb("ing") will(mod :neg), Adv?, negWord(type :post), vb(lemme :have), vb(partPast) will(mod :neg), Adv?, negWord(type :post), be(mod :aff), Adv?, vb("ing") will(mod :neg), Adv?, negWord(type :post), vb vb(lemma :have), Adv?, negWord(type :post), "been", vb("ing") vb(lemma :have), Adv?, negWord(type :post), vbOpinion(partPast) be(mod :neg), Adv?, vb("ing") negWord(type :pre), vb</p>	<p>"shouldn't eat" "will not have been eating" "will not have eaten" "will not be eating" "will not eat" "have not been eating" "haven't eaten" "am not only eating" "don't eat"</p>
	<p>SV(type :vb, mod :aff)</p>	<p>→ → → → → → → →</p>	<p>will, Adv?, vb(lemme :have), "been", vb("ing") will, Adv?, vb(lemme :have), vb(partPast) will, be(mod :aff), Adv?, vb("ing") will, be(mod :aff), Adv?, vb vb(lemme :have), "been", vb("ing") vb(lemme :have), vb(partPast) be(mod :aff), Adv?, vb("ing") vb tag :vb</p>	<p>"should eat" "will have eaten" "will be eating" "will eat" "have been eating" "have eaten" "am only eating" "eat"</p>

TABLEAU F.5 – Grammaires traitant les syntagmes verbaux n'appartenant pas aux catégories attributif, opinion, like ou dislike traitées par les grammaires précédentes

SN.grf	SN(pol :neutral)	→	Det(negMark :true), SAdj(pol :[positive   negative])+, N(pol :[positive   negative])	<i>any beautiful fight</i>	
	SN(pol :neutral)	→	Det(negMark :true), SAdj(pol :neutral), N(pol :[positive   negative])	<i>no cubist masterpiece</i>	
	SN(pol :neutral)	→	Det(negMark :true), SAdj(pol :[positive   negative]), N(pol :neutral)	<i>no beautiful painting</i>	
	SN(pol :neutral)	→	Det(negMark :true), N(pol :[neutral   positive   negative])	<i>no masterpiece</i>	
	SN(pol :SAdjPol)	→	Det(negMark :false) ?, SAdj(pol :[positive   negative]), N(pol :[positive   negative])	<i>a beautiful fight</i>	
	SN(pol :NPol)	→	Det(negMark :false) ?, SAdj(pol :neutral), N(pol :Pos   Neg)	<i>no french hero</i>	
	SN(pol :SAdjPol)	→	Det(negMark :false) ?, SAdj(pol :[positive   negative]), N(pol :neutral)	<i>a beautiful painting, a bad day</i>	
	SN(pol :NPol)	→	Det(negMark :false) ?, N(pol :[neutral   positive   negative])	<i>a painting, a masterpiece, a coward</i>	
	N(pol :positive)	→	Noun (InWordnetAffect, pol :positive)	<i>masterpiece</i>	
	N(pol :negative)	→	Noun (InWordnetAffect, pol :negative)	<i>coward</i>	
	N(pol :neutral)	→	tag :noun	<i>painting</i>	
	N(pol :neutral)	→	tag :pro	<i>I, you, me, he, they, etc</i>	
	SAdj.grf	SAdj(Pol)	→	Adv ?, Adj(Pol)	<i>very beautiful</i>
		Adj(pol :positive)	→	Adj(InWordnetAffect, pol :positive)	<i>beautiful, brave, fun, etc</i>
Adj(pol :negative)		→	Adj(InWordnetAffect, pol :negative)	<i>bad, ugly, sad, etc.</i>	
Adj(pol :neutral)		→	Adj(pol :neutral)	<i>blue</i>	
det.grf	Det(negMark :false)	→	"a", Noun, "of"	<i>a lot of</i>	
	Det(negMark :false)	→	such, tag :det	<i>such a</i>	
	Det(negMark :false)	→	tag :det	<i>a, the, your, my, etc.</i>	
	Det(negMark :true)	→	"any"   "no"   "neither"   "nor"		

TABLEAU F.6 – Grammaires traitant les syntagmes nominaux (SN.grf), les syntagmes adjectivaux (SAdj.grf) et les déterminants (det.grf)

F.2 Grammaires de niveau phrastique

L&D(pol : Pol <sub>Groupe2</sub> ) source : user cible : Groupe1	→	SN <sub>Groupe1</sub> , SV(lemme :make, mod :aff), SN(tag :proP1), SAdj(pol : [negative   positive]) <sub>Groupe2</sub>	This movie makes me happy
L&D(pol : Inverse(Pol <sub>Groupe2</sub> )) source : user cible : Groupe1	→	SN <sub>Groupe1</sub> , SV(lemme :make, mod :neg), SN(tag :proP1), SAdj(pol : [negative   positive]) <sub>Groupe2</sub>	This movie doesn't make me happy
L&D(pol : Pol <sub>Groupe1</sub> ) source : user cible : Groupe2	→	"it   this   that", SV(lemme :make, mod :aff), SAdj(pol : [negative   positive]) <sub>Groupe1</sub> , INF <sub>Groupe2</sub>	It makes me happy to work today
L&D(pol : Inverse(Pol <sub>Groupe1</sub> )) source : user cible : Groupe2	→	"it   this   that", SV(lemme :make, mod :neg), SAdj(pol : [negative   positive]) <sub>Groupe1</sub> , INF <sub>Groupe2</sub>	It doesn't make me happy to work today

TABLEAU F.7 – Grammaire reconnaissant les structures de type <Target> <Make> <Affect> <Source>, This movie makes me feel happy

L&D(pol : <i>Pol<sub>Groupe1</sub></i> source :user cible : <i>Groupe3</i> )	→	SN(tag :proP1), SV(lemme :attributif, mod :aff), [SN(pol[negative positive]) SAdj(pol :[negative positive])] <i>Groupe1</i> , (INF PART SP) <i>Groupe2</i>	<i>I was happy to see him yesterday</i>
L&D(pol : <i>Inverse(Pol<sub>Groupe1</sub>)</i> source :user cible : <i>Groupe3</i> )	→	SN(tag :proP1), SV(lemme :attributif, mod :neg), [SN(pol[negative positive]) SAdj(pol :[negative positive])] <i>Groupe1</i> , (INF PART SP) <i>Groupe2</i>	<i>I wasn't happy to see him yesterday</i>
L&D(pol : <i>Pol<sub>Groupe2</sub></i> source :user cible : <i>Groupe1</i> )	→	SN <i>Groupe1</i> , SV(lemme :attributif, mod :aff), [SN(pol[negative positive]) SAdj(pol :[negative positive])] <i>Groupe2</i>	<i>This movie is so bad</i>
L&D(pol : <i>Inverse(Pol<sub>Groupe2</sub>)</i> source :user cible : <i>Groupe2</i> )	→	SN <i>Groupe1</i> , SV(lemme :attributif, mod :neg), [SN(pol[negative positive]) SAdj(pol :[negative positive])] <i>Groupe2</i>	<i>This artwork isn't valuable</i>

 TABLEAU F.8 – Grammaire reconnaissant les structures de type attributif, *This movie is beautiful*

L&D(pol : <i>Pol</i> <sub>Groupe2</sub> ) source :user cible : <i>Groupe3</i> )	→	"it   this   that", SV(lemme :attributif, mod :aff, [SN(pol[negative positive]) SAdj(pol :[negative positive])]) <sub>Groupe2</sub> , (SP INF PART) <sub>Groupe3</sub> )	<i>It's a good idea to go to Peru for vacation</i>
L&D(pol : <i>Inverse</i> ( <i>Pol</i> <sub>Groupe2</sub> )) source :user cible : <i>Groupe3</i> )	→	"it   this   that", SV(lemme :attributif, mod :neg, [SN(pol[negative positive]) SAdj(pol :[negative positive])]) <sub>Groupe2</sub> , (SP INF PART) <sub>Groupe3</sub> )	<i>It isn't a good idea to go to Peru for vacation</i>

TABLEAU F.9 – Grammaire reconnaissant les structures de type attributif, *It's not a good idea to eat that*

L&D(pol : <i>Pol</i> <sub>Groupe1</sub> ) source :user cible : <i>Groupe3</i> )	→	SN(tag :proP1), SV(lemme :have, mod :aff, [SN(pol[negative positive]) SAdj(pol :[negative positive])]) <sub>Groupe1</sub> )	<i>I had a good week</i>
L&D(pol : <i>Inverse</i> ( <i>Pol</i> <sub>Groupe1</sub> )) source :user cible : <i>Groupe3</i> )	→	SN(tag :proP1), SV(lemme :have, mod :neg, [SN(pol[negative positive]) SAdj(pol :[negative positive])]) <sub>Groupe1</sub> )	<i>I had not a good week</i>

TABLEAU F.10 – Grammaire reconnaissant les structures de type <Source> <Have> <Affect>, *I had a awful week*

L&D(pol : Pol <sub>Groupe1</sub> ) source : user cible : Groupe3	→	SN(tag : proP1), SV(type : L&D, mod : aff), (SN   SP   INF   PART) <sub>Groupe1</sub>	I like outdoor activities
L&D(pol : Inverse(Pol <sub>Groupe1</sub> )) source : user cible : Groupe3	→	SN(tag : proP1), SV(type : L&D, mod : neg), (SN   SP   INF   PART) <sub>Groupe1</sub>	I don't like outdoor activities
L&D(pol : (Pol <sub>Groupe1</sub> )) source : user cible : Groupe3	→	SN(tag : proP1), SV(type : L&D, mod : aff), (SN   SP   INF   PART) <sub>Groupe1</sub>	I don't like outdoor activities
L&D(pol : Inverse(Pol <sub>Groupe1</sub> )) source : user cible : Groupe3	→	SN(tag : proP1), SV(type : L&D, mod : neg), (SN   SP   INF   PART) <sub>Groupe1</sub>	I don't like outdoor activities

TABLEAU F.11 – Grammaire reconnaissant les structures de type &lt;Source&gt; &lt;Like&gt; &lt;Target&gt;, I don't like you

L&D(pol : Pol <sub>Groupe2</sub> ) source : user cible : Groupe3	→	SN(tag : pro1) <sub>Groupe1</sub> , SV(lemme : feel, mod : aff), SAdj(pol : [negative   positive]) <sub>Groupe2</sub> , "about   to", [SN   Inf] <sub>Groupe3</sub>	I feel bad about that
L&D(pol : Inverse(Pol <sub>Groupe2</sub> )) source : user cible : Groupe3	→	SN(tag : proP1), SV(lemme : feel, mod : neg), SAdj(pol : [negative   positive]) <sub>Groupe2</sub> , "about   to", [SN   Inf] <sub>GroupeCible</sub>	I don't feel happy about my life

TABLEAU F.12 – Grammaire reconnaissant les structures de type &lt;Source&gt; &lt;Feel&gt; &lt;Target&gt;, I feel happy about life

L&D(pol : Pol <sub>Groupe2</sub> ) source : user cible : Groupe1	→	SN(tag : proP1), SV(type : opinion, mod : aff), (SN   PART) <sub>Groupe1</sub> , ("as   that") ?, [SN(pol [negative   positive])   SAdj(pol : [negative   positive])] <sub>Groupe2</sub>	<i>I consider this painting as beautiful</i>
L&D(pol : Inverse(Pol <sub>Groupe2</sub> )) source : user cible : Groupe1	→	SN(tag : proP1), SV(type : opinion, mod : neg), (SN   PART) <sub>Groupe1</sub> , [SN(pol [negative   positive])   SAdj(pol : [negative   positive])] <sub>Groupe2</sub>	<i>I don't consider this painting as beautiful</i>
L&D(pol : Pol <sub>Groupe1</sub> ) source : user cible : Groupe2	→	SN(tag : proP1), SV(type : opinion, mod : aff), ("it   that   this") ?, ("as   that") ?, [SN(pol [negative   positive])   SAdj(pol : [negative   positive])] <sub>Groupe1</sub> , (INF   PART) <sub>Groupe2</sub>	<i>I consider it as nice to help him</i>
L&D(pol : Inverse(Pol <sub>Groupe1</sub> )) source : user cible : Groupe2	→	SN(tag : proP1), SV(type : opinion, mod : neg), ("it   that   this") ?, ("as   that") ?, [SN(pol [negative   positive])   SAdj(pol : [negative   positive])] <sub>Groupe1</sub> , (INF   PART) <sub>Groupe2</sub>	<i>I don't consider it as nice to frighten him</i>

TABLEAU F.13 – Grammaire reconnaissant les structures à verbes d'opinion, *I think it's a good idea*





# Analyse des sentiments dans les conversations humain-agent. Vers un modèle des goûts de l'utilisateur

Caroline LANGLET

**RESUME** : Située à la croisée de deux domaines de recherche, celui de l'analyse de sentiments et celui des agents conversationnels animés, cette thèse s'intéresse à l'analyse du contenu verbal produit par l'utilisateur et à la manière dont celui-ci réfère ou exprime des sentiments, des affects ou des attitudes. Elle propose un modèle de détection de ces expressions au cours d'une interaction multi-modale et en face à face avec un agent conversationnel animé. Sur la base d'une analyse approfondie, un certain nombre de caractéristiques syntaxiques, sémantiques et pragmatiques sont mises en évidence. Leur formalisation permet le développement d'un modèle de détection symbolique intégrant des règles sémantiques et des grammaires formelles. Grâce à la création de deux plateformes d'annotation, deux jeux d'annotations ont été constitués. Les performances du modèle ont ainsi été évaluées par rapport aux références obtenues.

**MOTS-CLEFS** : Analyse de sentiments, Interactions humain-agent, Agents conversationnels animés.

**ABSTRACT** : This thesis is related to two research areas : sentiment analysis and embodied conversational agents. It focuses on the user's verbal expressions of sentiment and the development of a detection model for face-to-face human-agent conversations. First, an in-depth analysis of sentiment highlights a number of syntactic, semantic and pragmatic features. Then, we propose a symbolic model integrating semantic rules and formal grammars. Two annotation sets were then used for evaluating the performances of the model.

**KEY-WORDS** : Sentiment Analysis, Human-Agent Interactions, Embodied Conversational Agents.



