



HAL
open science

Prédire automatiquement les intentions du locuteur dans des questions issues du discours oral spontané

Angèle Barbedette, Iris Eshkol-Taravella

► To cite this version:

Angèle Barbedette, Iris Eshkol-Taravella. Prédire automatiquement les intentions du locuteur dans des questions issues du discours oral spontané. 6e conférence conjointe Journées d'Études sur la Parole (JEP, 31e édition), Traitement Automatique des Langues Naturelles (TALN, 27e édition), Rencontre des Étudiants Chercheurs en Informatique pour le Traitement Automatique des Langues (RÉCITAL, 22e édition), 2020, Nancy, France. pp.137-145. hal-02784761v1

HAL Id: hal-02784761

<https://hal.science/hal-02784761v1>

Submitted on 7 Jun 2020 (v1), last revised 23 Jun 2020 (v3)

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.



Distributed under a Creative Commons Attribution - NonCommercial - NoDerivatives 4.0 International License

Prédire automatiquement les intentions du locuteur dans des questions issues du discours oral spontané

Angèle Barbedette¹ Iris Eshkol-Taravella¹

(1) Université Paris Nanterre, MoDyCo UMR 7114, 200 avenue de la République, 92000 Nanterre, France
angele.barbedette@gmail.com, ieshkolt@parisnanterre.fr

RÉSUMÉ

Cette étude porte sur la classification automatique des intentions exprimées dans des questions issues d'un corpus d'échanges oraux spontanés. Nous proposons une typologie dans laquelle nous distinguons trois classes d'intentions (AVIS, VOLONTÉ et DOUTE). Après plusieurs prétraitements et ajouts de traits lexicaux aux données (lexiques, nombre de mots et de caractères), nous implémentons un algorithme de classification automatique et nous en présentons et évaluons les résultats qui atteignent une F-mesure de 0,62. Nous proposons ensuite une interprétation de ceux-ci, basée sur une comparaison entre les expériences menées et des mesures liées aux traits linguistiques intégrés avant la tâche de classification.

ABSTRACT

Automatically predicting the speaker's intentions in questions from spontaneous oral speech

This study focuses on the automatic classification of intentions expressed in questions from a corpus of spontaneous oral interactions. We suggest a typology in which we distinguish three categories of intentions (OPINION, WILL and DOUBT). After several preprocessings and additions of lexical features to the data (lexicons, number of words and characters), we implement an automatic classification algorithm and we present and evaluate the results which reach a F-measure of 0.62. We then provide an interpretation of these, based on a comparison between our experiments and some measures related to the linguistic features we integrated before the classification task.

MOTS-CLÉS : intentions, implicite, actes de dialogue, discours oral spontané.

KEYWORDS: intentions, implicit, dialog acts, spontaneous oral speech.

1 Introduction

L'objectif de ce travail est d'étudier les spécificités linguistiques de l'aspect non-littéral, c'est-à-dire de ce qui n'est pas dit de façon explicite, dans des questions issues de transcriptions d'échanges oraux et spontanés. Il s'agit d'une tâche primordiale pour tendre vers des systèmes de dialogue homme-machine plus performants et plus proches de conversations dites « naturelles ». Pour rendre compte de sa complexité, nous proposons un bref état de l'art des travaux en lien avec cette étude réalisés dans les domaines de la linguistique et du TAL.

L'énoncé est porteur d'une valeur illocutoire correspondant à la force ou l'intention qui lui est attribuée (Austin, 1962; Kerbrat-Orecchioni, 1986). Searle introduit de son côté la notion d'acte de langage indirect qui se compose d'un acte illocutoire primaire (non littéral), et d'un acte illocutoire

secondaire (littéral) (Searle, 1975). Dans (Allen & Perrault, 1980), les auteurs parlent également d'action intentionnelle pour définir l'acte de langage. Ces différents travaux introduisent la notion d'intention. Du point de vue du TAL, les travaux sur les intentions s'inscrivent dans les domaines de l'analyse d'opinion et des actes de dialogue. D'après (Karoui *et al.*, 2019, 2014), une opinion est une « expression subjective du langage » qui est soit explicite (repérée à l'aide d'indices textuels), soit implicite (s'appuyant sur des connaissances culturelles ou pragmatiques communes). Les actes de dialogue ont quant à eux pour objectif d'aider à l'analyse fine de la conversation et de tous les types d'énoncés qui la composent. Ils correspondent, selon Bunt, à la combinaison entre le contenu sémantique de l'énoncé et sa fonction communicative (Bunt, 2005). Plusieurs taxonomies d'actes de dialogue ont été proposées pour la tâche de classification automatique en actes de langage (Moldovan *et al.*, 2011) tel que le schéma d'annotation DIT ++ (Bunt, 2009) qui s'inspire d'autres taxonomies telles que DAMSL (Allen & Core, 1997), SWBD-DAMSL (Jurafsky *et al.*, 1997), HCRC Map Task (Anderson *et al.*, 1991) ou encore VERBMOBIL (Alexandersson *et al.*, 2011). Le dialogue est souvent étudié en TAL du point de vue de la sémantique formelle, avec des approches telles que celle de Ginzburg (Ginzburg, 1996a,b) et des théories telles que la SDRT par exemple (Segmented Discourse Representation Theory) (Maudet *et al.*, 2004), où l'identification de l'incompréhension entre les locuteurs s'appuie sur la recherche d'incohérences logiques dans les combinaisons entre les actes de parole, calculées à partir des algorithmes compositionnels fondés sur une représentation logique et dynamique d'une question et d'une réponse (Boritchev & Amblard, 2018, 2019). Selon ces théories, les nouveaux énoncés produits réalisent souvent des actes qui doivent être en lien avec d'autres éléments du contexte et s'inscrivent donc dans une structure complexe et non pas linéaire. Elles mettent en évidence l'importance de la prise en compte du contexte pour permettre la cohérence du dialogue.

Nous considérons, d'après la définition de l'énoncé performatif donnée par Austin (Austin, 1962) qu'en posant une question, le locuteur veut toujours dire quelque chose de plus que ce que la valeur locutoire de la question exprime en réalité. Les questions accomplissent toujours, selon nous et tels que les définit Searle, un acte illocutoire primaire et un acte illocutoire secondaire. L'association des caractéristiques des opinions et des actes de dialogue nous amène à définir ce que nous appelons dans cette étude l'*intention*. Il s'agit de l'activité illocutoire exprimée par un énoncé (« l'ensemble des actes qui s'accomplissent, immédiatement et spécifiquement, par l'exercice de la parole » (Ducrot, 1972)), qui permet de le caractériser selon son but, explicite ou implicite. L'activité illocutoire ne se restreint pas seulement à l'expression des opinions mais à tous les types d'objectifs impliqués par la production d'un énoncé. Dans (Chen *et al.*, 2013), les auteurs différencient les intentions explicites des intentions implicites. Des exemples illustrent ces deux cas : « I am looking for a brand new car to replace my old Ford Focus » qui correspond à une intention explicite d'achat et « Anyone knows the battery life of iPhone ? » qui est une intention implicite d'achat.

Ce travail s'intéresse particulièrement aux intentions implicites exprimées par les locuteurs lorsqu'ils posent une question et tente d'en effectuer la classification automatique ainsi que d'en dégager des spécificités linguistiques en se concentrant uniquement sur l'écrit. Il s'agit de tenir compte du contexte de chaque question à la fois pour l'étape des annotations manuelles et pour celle des annotations automatiques, sans se placer du point de vue de la linguistique formelle. Nous utilisons des techniques existantes ayant déjà fait leurs preuves dans la fouille d'opinion, un des domaines se rapprochant le plus de notre tâche.

2 Corpus, typologie et annotations

Les données utilisées proviennent des corpus oraux ESLO1 et ESLO2 (Enquêtes SocioLinguistiques à Orléans) (Baude & Dugua, 2011; Eshkol-Taravella *et al.*, 2011), comportant respectivement environ 300 et 400 heures d'enregistrement. Les transcriptions des enregistrements sont disponibles sur le site web d'ESLO¹ et ouvertes au public. Nous utiliserons ici toutes les transcriptions d'enregistrements effectués au cours de repas dans ESLO1 et ESLO2, soit un total de 28 fichiers au format *.xml* (sept issus du corpus ESLO1 et 21 issus d'ESLO2). Ils forment un tout d'environ 19 heures d'enregistrement. Le choix de cette catégorie précise est lié au fait de vouloir utiliser les données les plus spontanées possibles pour répondre à l'objectif principal de ce travail consistant à prédire l'intention du locuteur à travers les questions.

Après de premières observations pour déterminer des points communs et divergences entre les questions du corpus (ou cibles), nous avons fait plusieurs essais d'annotation, en nous demandant à chaque fois ce que le locuteur cherchait à exprimer à travers sa question. Ces essais d'annotation étaient entrecoupés de phases de discussion pour affiner les étiquettes, durant lesquelles se sont posées les questions de leur généricité et de leur applicabilité à l'ensemble du corpus. Un des objectifs était de parvenir à dégager des critères distinctifs pour chacune d'elles. Ces étapes nous ont permis de construire une typologie des intentions dans les questions se divisant en deux parties : la première s'intéressant au type de réponse attendu pour chacune des questions et donc à leur aspect explicite, et la seconde portant sur l'intention exprimée par le locuteur à travers sa question et donc plutôt sur son aspect implicite.

Pour l'explicite, nous avons défini deux classes permettant de catégoriser les questions en fonction du type de réponse attendue. Une question peut donc être :

- une *demande d'accord* (une interrogation totale), par exemple « je peux mettre ça là ? » ;
- une *demande d'information* (une interrogation partielle), par exemple « c'est où Saint Raphaël ? ».

L'implicite suppose des classes plus complexes à déterminer puisqu'elles nécessitent une interprétation : il s'agit d'un message non littéral. Nous avons dégagé trois classes représentant l'intention exprimée par le locuteur produisant la question (*avis*, *volonté* et *doute*), sur lesquelles nous nous concentrerons pour la tâche de classification automatique qui constitue la suite de ce travail :

- l'expression d'un *avis* correspond à un jugement positif ou négatif qui n'implique pas nécessairement une action de la part d'un des locuteurs (« ils sont vraiment bêtes hein ? ») et dont les marqueurs possibles peuvent être des adjectifs, des adverbes ou des locutions verbales qui aident à l'expression d'opinions (« je trouve que », « j'adore », « ennuyeux », « honnêtement », etc.) ;
- l'expression d'une *volonté* correspond au désir d'une action ou d'un comportement de la part du locuteur lui-même ou de son interlocuteur, implique une réponse correspondant à une action (« tu nous sers à boire mon chéri ? ») et peut avoir des marqueurs tels que des verbes d'action (« aller », « manger », « dormir », « regarder », etc.) ou des verbes exprimant une volonté (« vouloir », « souhaiter », « désirer », etc.) ;
- l'expression d'un *doute* correspond à une mise en doute de ce qui est dit, du caractère vrai ou faux d'une chose, qui n'implique pas nécessairement une action, qui s'apparente à une répétition, à de la surprise, à une demande de confirmation ou de précisions (« sûrement

1. <http://eslo.huma-num.fr>

non ? ») et dont les marqueurs possibles peuvent être la répétition d’une partie du contexte précédent, la présence de mots interrogatifs (« qui », « quel », « quoi », etc.) ou encore la présence d’adverbes d’affirmation/modaux (« sûrement », « peut-être », « probablement », etc.).

Nous avons utilisé les définitions de cette typologie pour annoter les 3647 questions de notre corpus, la prise de décision finale pour chacune d’entre elles s’appuyant également sur son contexte, c’est-à-dire sur les dix tours de parole la précédant et la suivant, celui-ci étant peu compréhensible avec une longueur inférieure. Les questions dites « phatiques » ont été considérées comme peu pertinentes pour cette tâche et n’ont donc pour la plupart pas été annotées.

Pour s’assurer de la fiabilité des définitions de nos classes et de nos annotations, nous avons mis en place un formulaire Google Forms pour l’annotation de quinze questions issues de notre corpus, associé à des consignes qui prévoyaient la prise en considération de leurs contextes. Ce formulaire a récolté vingt-six participations de locuteurs natifs du français, dont douze ayant une formation en linguistique. Nous avons ainsi calculé un accord inter-annotateur avec un Kappa de Cohen entre nos propres annotations et celles de chacun des participants. Nous avons obtenu pour 50% des participations un accord inter-annotateur supérieur à 0,73 pour l’explicite et supérieur à 0,6 pour l’implicite, et pour 25% des participants un accord supérieur à 0,86 pour l’explicite et supérieur à 0,8 pour l’implicite (Landis & Koch, 1977). Ces scores sont plutôt satisfaisants compte tenu de la nature de la tâche demandée qui nécessitait de s’appuyer à la fois sur des éléments concrets tels que l’analyse du contexte des questions et la prise en considération des indices lexicaux présents, mais aussi sur sa propre intuition.

3 Prétraitements et intégration de traits linguistiques

Pour préparer l’étape de classification automatique, nous avons procédé à la lemmatisation et à l’étiquetage morphosyntaxique des données à l’aide de *TreeTagger* (Schmid, 1994), et plus particulièrement à l’aide des fichiers de paramètres du projet *PERCEO, un Projet d’Étiqueteur Robuste pour l’Écrit et pour l’Oral* (Benzitoun et al., 2012), plus adaptés aux données orales et donc aux transcriptions d’enregistrements. Nous avons ensuite vectorisé avec un TF-IDF les questions et leurs contextes pour obtenir une représentation du texte utilisable en entrée de l’algorithme de classification. En comparaison d’autres types de vectorisations tels que *word2vec* avec les modèles *CBOW* et *Skip-Gram* ou les vecteurs préentraînés de *Flair*, le TD-IDF nous a permis d’obtenir des résultats de classification légèrement meilleurs.

La mise en place de notre typologie et les phases d’annotation nous ont fait remarquer des informations lexicales permettant de distinguer nos différentes classes. Elles constituent donc des indices que nous avons regroupés sous la forme de six lexiques, dont cinq construits spécifiquement pour ce travail :

- les *verbes de parole* : *dire, demander, proposer, suggérer, expliquer*, etc. ;
- les *verbes de mouvement*, issus de la ressource lexicale *DinaVmouv*, (Stosic & Aurnague, 2017) : *accrocher, suivre, s’asseoir, remplir, parcourir*, etc. ;
- les *mots interrogatifs* : *qui, combien, lequel, pourquoi, quand*, etc. ;
- les *interjections*, correspondant à la liste d’interjections présentées dans le guide de transcription des corpus ESLO : *mouais, hein, miam, bof, bah*, etc. ;
- les *sentiments* : *apprécier, ravi, haine, nul, inquiéter*, etc. ;

— les *adverbes et adjectifs modaux* : *vraiment, impossible, certainement, peut-être, vrai, etc.*

En plus des informations liées aux lexiques (la fréquence d’apparition des mots de chacun des lexiques dans chaque cible et contexte gauche ou droit), nous avons intégré à nos données le nombre de mots et le nombre de caractères par cible et par contexte.

4 Classification automatique : expériences et résultats

Après avoir équilibré les classes pour éviter un biais dans les résultats (286 occurrences par classe soit un total de 858 occurrences), nous avons implémenté l’algorithme *Random Forest* et utilisé la méthode de validation croisée *k-fold cross-validation* ainsi que la fonction *GridSearchCV* de *sklearn* pour choisir les valeurs optimales pour des hyperparamètres donnés, dans notre cas *n_estimators*, *criterion* et *bootstrap*. La performance du modèle a ensuite été évaluée avec une moyenne des mesures de précision, rappel et F-mesure obtenues pour chaque groupe.

		EXPÉRIENCES								
		1	2	3	4	5	6	7	8	9
TRAITS	vecteur c.	×	×	×	×	×	×	×	×	×
	vecteur g.		×							
	vecteur d.		×							
	POS tagging c.			×	×	×	×	×	×	
	POS tagging g.									
	POS tagging d.									
	sentiments c.				×	×	×	×	×	×
	sentiments g.					×				
	sentiments d.					×				
	interjections c.				×	×	×	×	×	×
	interjections g.					×				
	interjections d.					×				
	interrogatifs c.				×	×	×	×	×	×
	interrogatifs g.					×				
	interrogatifs d.					×				
	mouvement c.				×	×	×	×	×	×
	mouvement g.					×				
	mouvement d.					×				
	parole c.				×	×	×	×	×	×
	parole g.					×				
	parole d.					×				
	modaux c.				×	×	×	×	×	×
	modaux g.					×				
	modaux d.					×				
	explicité						×	×	×	×
	nb. mots c.							×	×	×
	nb. mots g.								×	×
	nb. mots d.								×	×
nb. caractères c.							×	×	×	
nb. caractères g.								×	×	
nb. caractères d.								×	×	
RÉSULTATS	Précision	0,622	0,492	0,618	0,632	0,6	0,63	0,621	0,631	0,613
	Rappel	0,612	0,493	0,612	0,623	0,592	0,622	0,617	0,624	0,606
	F-mesure	0,611	0,489	0,61	0,622	0,59	0,62	0,616	0,622	0,603

TABLE 1 – Récapitulatif des expériences et résultats

Nous avons testé l’algorithme avec plusieurs combinaisons de traits présentées dans le tableau 1, un récapitulatif global de l’ensemble des résultats. Chaque expérience correspond à un ensemble choisi de traits (cochés dans le tableau). Nous voyons par exemple que toutes les expériences incluent la vectorisation de la cible mais que seules les expériences 8 et 9 incluent le nombre de caractères pour les contextes gauche et droit. Les résultats sont très proches dans l’ensemble comme nous le voyons pour les expériences 4, 6 et 8 pour lesquelles nous obtenons des scores de F-mesure avoisinant 0,62. Certaines mesures semblent cependant se détacher, telles que celles de l’expérience 2 qui sont inférieures à 0,5 ou celles de l’expérience 5 inférieures à 0,6, ces deux expériences prenant en compte des traits liés aux contextes de la question cible (la vectorisation dans le premier cas et la présence de mots des lexiques dans le second).

5 Discussion

Une première observation concerne la baisse des performances lorsque des traits liés aux contextes sont ajoutés, comme dans les expériences 2 et 5. Pour vérifier l'importance du contexte des cibles, nous avons reproduit les expériences 2 et 8 avec deux puis cinq tours de parole avant et après la question. Les scores obtenus pour l'expérience 2 (qui ne comprenait que la vectorisation des cibles et de leurs contextes) montrent une amélioration des performances lorsqu'il y a moins de tours de parole dans chaque contexte, ce qui peut s'expliquer par la trop grande quantité d'informations non pertinentes rapportées par la vectorisation des contextes lorsque ceux-ci sont plus larges. En revanche, nous observons pour l'expérience 8 (qui comprenait la vectorisation des cibles et des informations lexicales) que les scores sont meilleurs lorsque la fenêtre de contexte est plus grande : lorsque la vectorisation des contextes est utilisée en entrée du classifieur, elle semble rapporter beaucoup de bruits et faire baisser les performances, tandis que lorsque l'algorithme s'appuie sur des informations lexicales plutôt que sur la vectorisation, il est plus performant avec plus de contextes.

Nb tours de parole par contexte	Expérience 2			Expérience 8		
	2	5	10	2	5	10
Précision	0,575	0,54	0,492	0,588	0,591	0,631
Rappel	0,572	0,536	0,493	0,584	0,587	0,624
F-mesure	0,571	0,535	0,489	0,582	0,585	0,622

TABLE 2 – Comparaison des expériences 2 et 8 en fonction du nombre de tours de parole par contexte

Pour interpréter les résultats, nous nous sommes concentrés sur l'expérience 8, une des expériences ayant eu les meilleurs scores et prenant en compte le plus de traits. Nous avons calculé la médiane du nombre de mots et du nombre de caractères pour chaque cible, contexte gauche et contexte droit. Nous avons également calculé le nombre de mots appartenant aux différents lexiques présents en moyenne dans chaque cible (tableau 3).

	A→A	A→D	A→V	V→V	V→A	V→D	D→D	D→A	D→V
méd. mots c.	6	5	5	6	5	5	3	5	6
méd. mots g.	96	92	93,5	87,5	95	88	92	85	101
méd. mots d.	97	97	81	88,5	97	88	84	94	92
méd. caractères c.	21	19	21	21	19	17	11	17	21
méd. caractères g.	331	326	330,5	312,5	322	302	327	312	341
méd. caractères d.	344	338	301	310	345	317	299	331	309,5
vb. de parole	0,03	0,04	0,07	0,04	0,04	0,02	0,05	0,08	0,07
vb. de mouvement	0,06	0,07	0,12	0,19	0,12	0,1	0,08	0,11	0,16
mots interrogatifs	0,34	0,53	0,47	0,33	0,35	0,37	0,58	0,46	0,59
interjections	0,54	0,27	0,25	0,18	0,37	0,3	0,08	0,3	0,24
sentiments	0,41	0,09	0,12	0,03	0,09	0,13	0,04	0,16	0,1
modaux	0,31	0,07	0,09	0,13	0,26	0,13	0,09	0,11	0,07

TABLE 3 – Mesures pour les prédictions des classes avis (A), volonté (V) et doute (D)

Pour la classe *doute*, la médiane pour le nombre de caractères des questions est de 11. Il s'agit d'une longueur courte par rapport aux questions de la classe *doute* ayant été mal classées en *avis* ou en *volonté* pour lesquelles nous trouvons respectivement des médianes de 17 et de 21. Pour cette classe, la présence de mots interrogatifs est également significative puisqu'elle est plus élevée en moyenne qu'ailleurs (0,58). Les questions de cette classe ayant été mal classées en *volonté* contiennent en moyenne 0,16 verbes de mouvement, un chiffre qui se rapproche de la moyenne du nombre de verbes de mouvement présents dans les bonnes prédictions de *volonté* (0,19).

Pour la classe *avis*, les cibles bien classées ont en moyenne une forte présence de mots appartenant aux lexiques des sentiments (0,41), des interjections (0,55) et des modaux (0,31) comparé aux autres classes et aux cibles mal classées de la classe *avis*. Ces dernières, lorsqu'elles sont mal classées dans *volonté* ont en moyenne 0,12 mots appartenant au lexique des sentiments, 0,25 mots appartenant aux interjections et 0,09 mots étant des modaux. Ceci se confirme également pour les cibles *avis* mal classées en *doute*, pour lesquelles nous voyons par ailleurs que le nombre de mots interrogatifs présents en moyenne est de 0,53, un chiffre qui se rapproche de celui obtenu pour les cibles *doute* bien classées qui est de 0,58.

Enfin, pour la classe *volonté*, nous remarquons une plus forte présence de verbes de mouvement qu'ailleurs, avec une moyenne de 0,19 pour les cibles bien classées en *volonté*. Ce chiffre est plus bas pour les cibles mal classées en *doute* (0,1) et se rapproche notamment de la moyenne du nombre de verbes de mouvement pour les bonnes prédictions de *doute* qui est de 0,08. Pour les cibles de *volonté* mal classées en *avis*, nous observons une présence plus forte de mots des lexiques d'interjections (0,37) et de modaux (0,26), qui sont des caractéristiques de la classe *avis*.

6 Perspectives et conclusions

Le but de cette étude était de parvenir à définir et à identifier les intentions implicites exprimées par les locuteurs. Il s'agissait plus précisément d'étudier les spécificités linguistiques permettant de caractériser l'aspect non-littéral de questions issues du discours oral spontané, en prenant en considération leurs contextes. Après avoir constitué un corpus de référence dont nous avons évalué les étiquettes créées à partir de nos recherches et de notre typologie, nous avons procédé à une tâche de classification automatique avec Random Forest pour laquelle nous avons également intégré des traits linguistiques à nos données originales. Les scores obtenus pour ces expériences sont supérieurs au hasard mais restent tout de même améliorables puisqu'ils tournent pour la plupart autour de 0,6. Ces résultats semblent constituer un plafond et rendent ainsi compte de la complexité de la tâche effectuée. Ils restent cependant encourageants lorsqu'ils sont mis en regard avec les accords inter-annotateurs obtenus lors de l'évaluation des annotations manuelles, bons ou excellents pour la moitié d'entre eux.

Plusieurs idées et perspectives pourraient améliorer ces résultats :

- l'augmentation de la taille du corpus de référence pour permettre un meilleur apprentissage ;
- l'identification de nouveaux traits discriminants ;
- l'enrichissement de traits existants (par exemple l'ajout de nouvelles entrées à nos lexiques) ;
- l'utilisation de techniques d'apprentissage profond.

Une perspective à plus long terme de ce travail serait d'adapter cette étude sur les intentions implicites à d'autres types d'énoncés que les questions ou encore à d'autres types de données que les transcriptions de discours oral et spontané.

Références

ALEXANDERSSON J., BUSCHBECK-WOLF B., FUJINAMI T., KIPP M., KOCH S., MAIER E., REITHINGER N., SCHMITZ B. & SIEGEL M. (2011). *Dialogue acts in VERBMOBIL-2*, volume 1998. 2. ed. édition. [Report 226](#).

- ALLEN J. & CORE M. (1997). Draft of DAMSL : Dialog Act Markup in Several Layers. <https://www.cs.rochester.edu/research/speech/damsl/RevisedManual/>.
- ALLEN J. F. & PERRAULT C. R. (1980). Analyzing intention in utterances. *Artificial intelligence*, **15**(3), 143–178.
- ANDERSON A. H., BADER M., BARD E. G., BOYLE E., DOHERTY G., GARROD S., ISARD S., KOWTKO J., MCALLISTER J., MILLER J. *et al.* (1991). The HCRC map task corpus. *Language and speech*, **34**(4), 351–366.
- AUSTIN J. L. (1962). *How to do things with words*. William James Lectures. Oxford University Press.
- BAUDE O. & DUGUA C. (2011). (Re)faire le corpus d’Orléans quarante ans après : quoi de neuf, linguiste ? *Corpus*, **10**, 99–118.
- BENZITOUN C., FORT K. & SAGOT B. (2012). TCOF-POS : un corpus libre de français parlé annoté en morphosyntaxe. In *Actes de la conférence conjointe JEP-TALN-RECITAL 2012*. HAL : [hal-00709187](https://hal.archives-ouvertes.fr/hal-00709187).
- BORITCHEV M. & AMBLARD M. (2018). Coffee or tea ? Yes. In L. PRÉVOT, M. OCHS & B. FAVRE, Édts., *Proceedings of the 22nd workshop on the Semantics and Pragmatics of Dialogue*. HAL : [hal-01922137](https://hal.archives-ouvertes.fr/hal-01922137).
- BORITCHEV M. & AMBLARD M. (2019). A compositional view of questions. In *Proceedings of the 2019 Workshop on Widening NLP*. HAL : [hal-02269603](https://hal.archives-ouvertes.fr/hal-02269603).
- BUNT H. (2005). A framework for dialogue act specification. *Proceedings of SIGSEM WG on Representation of Multimodal Semantic Information*.
- BUNT H. (2009). The DIT++ taxonomy for functional dialogue markup. In *AAMAS 2009 Workshop, Towards a Standard Markup Language for Embodied Dialogue Acts (EDAML 2009)*, p. 13–24.
- CHEN Z., LIU B., HSU M., CASTELLANOS M. & GHOSH R. (2013). Identifying intention posts in discussion forums. In *Proceedings of the 2013 conference of the North American chapter of the association for computational linguistics : human language technologies*, p. 1041–1050.
- DUCROT O. (1972). *Dire et ne pas dire : principes de sémantique linguistique*. Savoir. Hermann.
- ESHKOL-TARAVELLA I., BAUDE O., MAUREL D., HRIBA L., DUGUA C. & TELLIER I. (2011). Un grand corpus oral « disponible » : le corpus d’Orléans 1 1968-2012. *Traitement Automatique des Langues*, **52**(3). HAL : [halshs-01163053](https://hal.archives-ouvertes.fr/halshs-01163053).
- GINZBURG J. (1996a). Dynamics and the semantics of dialogue. In J. SELIGMAN & D. WESTERSTÅHL, Édts., *Logic, Language and Computation, Volume 1*, volume 58 de *CSLI Lecture Notes*, chapitre 15. CSLI Publications.
- GINZBURG J. (1996b). Interrogatives : Questions, facts and dialogue. In *The handbook of contemporary semantic theory*. Blackwell, Oxford, p. 359–423. Citeseer.
- JURAFSKY D., SHRIBERG E. & BIASCA D. (1997). *Switchboard SWBD-DAMSL Shallow-Discourse-Function Annotation Coders Manual, Draft 13*. Rapport interne 97-02, Institute of Cognitive Science University of Colorado, Boulder. <https://www.colorado.edu/ics/technical-reports-1990-1999>, Part 1, Part 2.
- KAROUI J., BENAMARA F. & MORICEAU V. (2019). *Détection automatique de l’ironie : Application à la fouille d’opinion dans les microblogs et les médias sociaux*. ISTE Group.
- KAROUI J., GILLES N. A., BENAMARA ZITOUNE F. & BELGUITH L. (2014). Le langage figuratif dans le web social : cas de l’ironie et du sarcasme. In *Workshop Fouille d’opinion dans le Web social*, Lyon, France. HAL : [hal-01686491](https://hal.archives-ouvertes.fr/hal-01686491).

- KERBRAT-ORECCHIONI C. (1986). *L'implicite*. Paris : A. Colin.
- LANDIS J. R. & KOCH G. G. (1977). The measurement of observer agreement for categorical data. *Biometrics*, **33**(1), 159–174.
- MAUDET N., MULLER P. & PRÉVOT L. (2004). Tableaux conversationnels en SDRT. *Actes de TALN*.
- MOLDOVAN C., RUS V. & GRAESSER A. (2011). Automated speech act classification for online chat. In S. VISA, A. INOUE & A. RALESCU, Édés., *Proceedings of The 22nd Midwest Artificial Intelligence and Cognitive Science Conference 2011*, p. 23–29. <http://ceur-ws.org/Vol-710/paper22.pdf>.
- SCHMID H. (1994). Probabilistic Part-of-Speech Tagging Using Decision Trees. In *Proceedings of the International Conference on New Methods in Language Processing*, Manchester, UK.
- SEARLE J. (1975). Indirect Speech Acts. In P. COLE & J. L. MORGAN, Édés., *Syntax and Semantics, Volume 3 : Speech Acts*. Academic Press.
- STOSIC D. & AURNAGUE M. (2017). DinaVmouv : Description, INventaire, Analyse des Verbes de mouvement. An annotated lexicon of motion verbs in French. Ressource lexicale, HAL : [hal-01979613](https://hal.archives-ouvertes.fr/hal-01979613).