

# Le langage figuratif dans le web social : cas de l'ironie et du sarcasme

Jihen Karoui<sup>12</sup>, Nathalie Aussenac Gilles<sup>3</sup>, Farah Benamara Zitoune<sup>2</sup>, Lamia Hadrich Belguith<sup>1</sup>

<sup>1</sup> MIRACL-Université de Sfax-Tunisie

<sup>2</sup> IRIT-Université de Toulouse,

<sup>3</sup> IRIT-CNRS, Toulouse

{Jihen.Karoui,benamara,Nathalie.Aussenac-Gilles}@irit.fr  
l.belguith@fsegs.rnu.tn

**Résumé.** L'analyse automatique du langage figuratif est l'un des défis majeurs du traitement des langues. Contrairement au langage littéral, le langage figuratif détourne le sens propre pour lui conférer un sens dit figuré ou imagé, comme la métaphore, l'ironie, le sarcasme, la satire et l'humour. La détection de ces phénomènes requiert des outils plus complexes que ceux utilisés pour l'analyse d'opinion. Dans ce contexte, nous nous focalisons sur l'ironie et le sarcasme. Nous présentons un panorama des définitions et principaux travaux existants. Nous décrivons ensuite une première expérimentation qui vise à détecter l'ironie dans un corpus de tweets en français.

**Mots clés:** Ironie, Sarcasme, langage figuratif.

## 1. Introduction

L'analyse des sentiments est un domaine de recherche extrêmement actif en traitement automatique des langues. En effet, ces dernières années ont vu se multiplier les sources de données textuelles porteuses d'opinion disponibles sur le web. Devant cette abondance de données et de sources, l'automatisation de la synthèse des multiples avis devient cruciale pour obtenir efficacement une vue d'ensemble des opinions sur un sujet donné. Globalement, les systèmes actuels ont obtenus de bons résultats sur la classification automatique du caractère subjectif ou objectif d'un document (Liu, 2010). En revanche, les résultats sur la tâche d'analyse de polarité restent encore peu concluants. La raison principale de cet échec est l'incapacité des algorithmes actuels à comprendre toutes les subtilités du langage humain. Parmi ces subtilités, nous focalisons dans cette communication sur le langage figuratif et plus particulièrement sur l'ironie et le sarcasme.

Contrairement au langage littéral, le langage figuratif détourne le sens propre pour lui conférer un sens dit figuré ou imagé. Selon le TLFi, l'ironie est une figure de rhétorique par laquelle on dit le contraire de ce qu'on veut faire comprendre. L'ironie recouvre un ensemble de phénomènes distincts dont les principaux sont l'ironie verbale et l'ironie situationnelle (Niogret, 2004). L'ironie verbale exprime une contradiction entre la pensée du locuteur et son expression. Elle est créée par le langage. L'ironie situationnelle désigne toute situation qui vient contredire les propos ou les prétentions d'une personne. Le sarcasme est généralement vu comme un sous-type de l'ironie verbale (Dews et Winner, 1995). Selon (Simedoh, 2012), le sarcasme est une ironie mordante parce qu'elle s'exprime avec aigreur et emportement. Elle s'exerce en présence de celui qu'il vise et souligne donc l'agressivité. Cette agressivité n'empêche pas que le sarcasme puisse comporter aussi des scènes de raillerie ou de moquerie.

Dans la suite, nous nous limitons à l'étude de l'ironie verbale en ne faisant aucune distinction entre ironie et sarcasme étant donné que la frontière entre ces deux notions est souvent floue (Reyes et al., 2013). Voici quelques exemples, extraits de commentaires web, qui illustrent ces phénomènes:

- « J'adore la façon dont votre produit tombe en panne dès que j'en ai besoin »,
- « Merci une fois de plus la SNCF. Ça annonce une magnifique journée ça encore »
- « Super match les bleus ! »,
- « Le camion de déménagement arrive déjà à l'Élysée ? »
- « la prison en Russie, c'est pas le club med »

Les trois premiers commentaires expriment des opinions négatives bien que les auteurs utilisent des mots d'opinion positifs (adorer, merci, magnifique, super). Les deux derniers expriment deux opinions négatives implicites (i.e. non lexicalisées). La détection automatique de ces phénomènes requiert donc des outils plus complexes que ceux utilisés classiquement pour l'analyse d'opinion.

Les travaux traitants de l'ironie et du sarcasme peuvent être divisés en deux catégories : les approches linguistiques qui s'intéressent à l'étude de ces phénomènes d'un point de vue langagier et philosophique et les approches computationnelles qui utilisent les techniques de TAL pour détecter automatiquement la présence de ces phénomènes dans le discours. Nous présentons un panorama des principaux travaux existants. Nous décrivons ensuite une première expérimentation qui vise à détecter l'ironie dans un corpus de tweets en français.

## 2. Etat de l'art

### 2.1 Les approches linguistiques

Deux grandes visions cohabitent : celle de (Grice, 1975) qui traite l'ironie verbale comme négation et celle de (Sperber et Wilson, 1981) qui la traite comme une variété d'utilisation échoïque de la langue. Les autres approches sont jugées, en général, du point de vue de leur proximité ou de leur éloignement théorique d'une des deux théories ci-dessus. Selon Grice, l'ironie consiste dans l'emploi d'un énoncé qui normalement signifiera « p » pour transmettre « non-p ». Il considère la description de l'ironie comme une violation la plus importante de la maxime de la qualité. En revanche (Sperber et Wilson, 1981) exploitent la distinction de l'emploi vs. la mention et décrivent l'ironie comme une forme spéciale de la mention du langage par laquelle un locuteur répète une proposition ou une pensée attribuée à quelqu'un d'autre. En répétant cette proposition, le locuteur fait connaître à l'interlocuteur son attitude critique vis-à-vis de son contenu.

Dans la lignée des travaux de Grice, citons enfin quelques autres théories dans ce domaine: la théorie du faux-semblant de (Clark et Gerrig, 1984), l'interprétation de la signification des énoncés ironique en termes de manifestation de sa saillance cognitive (Giora, 1995), la théorie du faux-semblant allusif de (Kumon-Nakamura et al., 1995) où l'ironie n'est plus le décalage entre le dit et le signifié mais entre « ce qui est dit » et « ce qui aurait dû être dit au regard du contexte », la théorie d'Utsumi (1996) qui consiste à unifier l'ironie dans lequel il définit formellement un énoncé ironique et enfin la théorie d'Attardo (2000) qui définit un énoncé ironique comme un énoncé inapproprié au regard du contexte, qui reste néanmoins pertinent dans l'interaction.

Les modèles théoriques décrits plus hauts sont à la base des approches computationnelles de détection de l'ironie et du sarcasme. En effet, la plupart des indices ou traits utilisés pour la détection automatique ont été inspirés par ces travaux.

### 2.2 Les approches computationnelles

La majorité de travaux en TAL se focalisent sur les textes d'opinions, comme des avis de consommateurs ou des textes courts issus de réseaux sociaux comme twitter. En général, les avis de consommateurs négatifs sont supposés avoir plus de chance de contenir des expressions ironiques (Tsur et al., 2010), ce qui est évidemment discutable. Pour le corpus de tweets, la collecte de messages associés aux #tags #sarcasme, #ironie, #humour, #satire sont considérés comme ironiques ou sarcastiques alors que des #tags comme #politique ou encore #éducation ne le sont pas (Reyes et al., 2013). Plusieurs corpus annotés en ironie sont actuellement librement disponibles. L'annotation consiste à associer à chaque document soit une catégorie binaire (ironique/non ironique) soit un score ou un degré d'ironie ou de sarcasme le long d'une échelle donnée. Parmi les ressources existantes, citons le corpus Senti-TUT<sup>1</sup> formé de tweets Italiens (Gianti et al., 2012), le corpus de (Burfoot et al., 2009) formé d'articles de presse satirique anglais<sup>2</sup> et le corpus de (Filatova, 2012) formé à partir des commentaires sur les produits d'Amazon<sup>3</sup>.

Globalement, trois principales approches ont été proposées pour détecter l'ironie: les méthodes psycholinguistiques, les méthodes symboliques et les méthodes d'apprentissage. La première approche permet de tester certaines hypothèses linguistiques sur l'ironie en les confrontant aux jugements d'annotateurs humains (via par exemple des plateformes de type Mechanical Turk). On présente aux annotateurs un ensemble de textes ou expressions, et les annotateurs doivent juger de leurs caractères ironiques ou non ironiques selon un ensemble de traits ou d'indices linguistiques. (Utsumi, 2004) a montré que

<sup>1</sup> <http://www.di.unito.it/~tutreeb/sentitut.html>

<sup>2</sup> <http://www.csse.unimelb.edu.au/research/lt/resources/satire/>

<sup>3</sup> <http://storm.cis.fordham.edu/~filatova/SarcasmCorpus.html>

l'opposition, les questions rhétoriques et le niveau de politesse sont des indices pertinents. (Kreuz et al., 2007) concluent que la présence d'interjection est important alors que le nombre de mots, la présence d'adjectifs et d'adverbes, de mots en gras ou encore la ponctuation n'a pas d'influence. Enfin, (Gianti et al., 2012) ont en plus identifié que le temps des verbes peut être une autre façon d'étudier les différences linguistiques entre les textes humoristiques et les textes objectifs.

Les secondes méthodes ont pour but de proposer des patrons lexico-syntaxiques afin de détecter le langage figuratif sans l'utilisation d'outils d'apprentissage. (Sarmiento et al., 2009) ont proposé un ensemble de 8 patrons d'extraction. Les meilleurs patrons sont ceux basés sur les citations et les émoticons avec une précision respective de 85,4% et 68,3%, sur un corpus d'articles de presse en portugais. (Vaelo et al., 2010) repèrent des expressions de similitudes ironiques qui répondent à des patrons comme « about as \* as \* » ou encore « as ADJ as » telles que « *As tough as a marshmallow cardigan* ». Ils concluent que les similitudes ironiques expriment souvent des sentiments négatifs en utilisant des termes positifs. L'évaluation de ce modèle obtient une F-mesure de 73% pour la détection de l'ironie et une F-mesure de 93% pour la détection des exemples non-ironiques.

Enfin, les méthodes d'apprentissage se focalisent essentiellement sur du semi-supervisé en utilisant à la fois des données étiquetées et non-étiquetées. (Tsur et al., 2010) proposent un algorithme appelé « SASI » pour la classification des phrases en sarcastique ou non sarcastique. Pour ce faire, ils ont définie deux types de traits à savoir des traits basés sur des patrons de mots les plus fréquents dans le corpus et des traits syntaxiques basés sur la ponctuation. Par la suite, ils ont appliqué ces traits sur un corpus formé de 60,000 commentaires sur les produits d'Amazon. La combinaison de tous les traits a donné une valeur de précision de 91,2% et une valeur de F-mesure de 82,7%. Les signes de ponctuation sont les prédicateurs les plus faibles. (Gonzalez-Ibanez et al., 2011) ont entrepris une étude empirique sur l'utilisation de traits lexicaux tels que les unigrams, les dictionnaires, les interjections et des traits pragmatiques (les émotions positives et négatives) pour distinguer entre le sarcasme et les sentiments positifs et négatifs exprimés dans les messages Twitter. (Reyes et al., 2011)(Reyes et al., 2012) ont proposé un ensemble de traits (n-grams, n-grams POS, profilage fanny, profilage positif/négatif, profilage affective, profilage d'agrément). Par la suite, ils ont étudié la pertinence de ces traits en regardant la possibilité de l'automatisation de l'identification des documents ironiques. Citons enfin quelques rares travaux basés sur de l'apprentissage supervisé qui utilisent que des données étiquetées. (Burfoot et al., 2009) proposent la combinaison d'un ensemble de traits lexicaux (titre, profanation, l'argot). Enfin, (Reyes et al., 2013) ont proposé un ensemble de traits comme les signatures (pointillisme, contre-factualité et compression temporelle), les scénarios émotionnels (activation, imagerie et plaisir) et l'inattendu (déséquilibre temporel et déséquilibre contextuelle). Les documents longs ont plus tendance à contenir du langage figuratif.

### 3. Une première expérimentation

La plupart des travaux de l'état de l'art se focalisent sur l'anglais. Quelques travaux ont été réalisés dans d'autres langues comme l'italien, le portugais ou le japonais. Cependant, à notre connaissance, aucun travail n'a abordé cette problématique pour le français.

En faisant un aperçus sur les différentes méthodes proposées, nous pouvons dégagés un ensemble de traits qui sont jugés discriminants (l'opposition, les questions rhétoriques, la circonlocution, le niveau de politesse (énoncés honorifiques), la présence d'interjection, la négation, le nombre des citations dans la phrase, le nombre des mots en majuscules dans la phrase) pour la détection de l'ironie. Comme on peut dégager les traits jugés non-discriminants (la présence des adjectifs et des adverbes, l'utilisation de la ponctuation) et les traits jugés non utiles (le nombre de mots, et le nombre de mots en caractère gras).

Nous proposons une première expérimentation qui vise à tester l'efficacité des traits de l'état de l'art sur la détection de messages ironiques extraits de Tweeter. Pour ce faire, nous proposons d'extraire un ensemble de 10,000 tweets relatifs aux #tags (#ironie, #sarcasme). Ensuite, nous nous focaliserons sur l'étude de trois types de traits : stylistiques (émoticons, la ponctuation, citations, mots en majuscule), lexicaux (n-grammes, adjectifs, adverbes, interjections), syntaxiques (POS) et sémantique (mots d'opinion, marqueurs de discours, etc.). Nous envisageons de tester l'efficacité de chaque trait en utilisant des techniques de sélections de traits telles que la fréquence d'apparition ou le gain d'information. Nous proposons ensuite un premier modèle de classification de tweets en deux classes ironiques et non ironiques. Les expérimentations sont actuellement en cours. Les résultats seront communiqués lors du workshop en cas d'acceptation de notre proposition.

## Bibliographie

- Attardo S. Irony as relevant inappropriateness. *Journal of Pragmatics* 32(6):793–826 (2000).
- Burfoot C., Baldwin T. Automatic Satire Detection: Are You Having a Laugh?. Proceedings of the ACL-IJCNLP 2009 Conference Short Papers, pp. 161-164 (2009).
- Clark H., Gerrig R. On the pretense theory of irony. *Journal of Experimental Psychology*, 113(1):121-6 (1984).
- Dews Sh., Winner E. Muting the meaning: A social function of irony. *Metaphor and Symbolic Activity*, 10(1):3–19 (1995).
- Gianti A., Bosco C., Patti V., Bolioli A., Di Caro L. Annotating Irony in a Novel Italian Corpus for Sentiment Analysis. In Proceedings of the 4th International Workshop on Corpora for Research on Emotion Sentiment & Social Signals (ES3) at LREC'12 (2012).
- Giora R. On irony and negation. *Discourse Processes*, 19(2):239–264 (1995).
- González-Ibáñez R., Muresan S., Wacholder N. Identifying sarcasm in Twitter: a closer look. In Proceedings of the 49th Annual Meeting of the Association for Computational Linguistics: Human Language Technologies: short papers - Volume 2 (HLT '11), Vol. 2. Association for Computational Linguistics, Stroudsburg, PA, USA, 581-586 (2011).
- Grice H.P. Logic and conversation. In Peter Cole and Jerry Morgan, editors, *Syntax and Semantics 3: Speech Acts*, pp. 64–75. Academic Press, New York (1975).
- Grice H.P. Logic and conversation, In P. Cole and J.L. Morgan (Eds.), *Syntax and semantics: Vol. 9. Pragmatics*. New York: Academic (1978).
- Grice, H.P. *Studies in the Way of Words*. Harvard University Press. (SWW) (1989).
- Kreuz R.J., Caucci G.M.. Lexical Influences on the Perception of Sarcasm. Proceedings of the Workshop on Computational Approaches to Figurative Language, pp. 1–4, Rochester, NY, Association for Computational Linguistics (2007).
- Kumon-Nakamura S., Glucksberg S. How about another piece of pie: the allusional pretense theory of discourse irony. *Journal of Experimental Psychology*, 124(1):3-21. (1995).
- Liu, Bing. Sentiment Analysis and Subjectivity. in N. Indurkha and F. J. Damerau (eds.), *Handbook of Natural Language Processing*, , CRC Press, 627–666 (2010).
- Niogret Ph. Les figures de l'ironie dans A La Recherche du Temps perdu de Marcel Proust. Harmattan (Editions L') Collection : Approches Littéraires (2004).
- Reyes A., Rosso P. Mining Subjective Knowledge from Customer Reviews: A Specific Case of Irony Detection. Proceedings of the 2nd Workshop on Computational Approaches to Subjectivity and Sentiment Analysis, ACL-HLT 2011, Portland, Oregon, USA, pages 118–124 (2011).
- Reyes A., Rosso P. Making objective decisions from subjective data: Detecting irony in customer reviews. *Decision Support Systems* volume 53 Issue 4: pp. 754-760 (2012).
- Reyes A., Rosso P., Vealo T. A multidimensional approach for detecting irony in Twitter. *Language Resources and Evaluation*, Volume 47, Issue 1, pp. 239-268 (2013).
- Sarmento L., Carvalho P., Silva M.J., de Oliveira E. Clues for Detecting Irony in User-Generated Contents: Oh...!! It's "so easy" ;-). TSA'09, Hong Kong, China (2009).
- Simedoh K.V. L'humour et l'ironie en littérature Francophone Subsaharienne. *Francophone Cultures and Literatures* - New York, Bern, Berlin, Bruxelles, Frankfurt am Main, Oxford, Wien, volume 60, pp. 180 (2012).
- Sperber D., Wilson D. Irony and the use-mention distinction. In P. Cole (ed.) *Radical pragmatics*. Academic Press, New York: 295-318. Reprinted in S. Davis (ed.) 1991 *Pragmatics: A reader*. Oxford University Press, Oxford: 550-63 (1981).
- Tsur O., Davidov D., Rappoport Ari. ICWSM – A Great Catchy Name: Semi-Supervised Recognition of Sarcastic Sentences in Online Product Reviews. Proceedings of the Fourth International AAAI Conference on Weblogs and Social Media (2010).

Utsumi A. A unified theory of irony and its computational formalization. In COLING '96 Proceedings of the 16th conference on Computational linguistics-volume 2, pp. 962-967, Association for Computational Linguistics, Stroudsburg (PA) (1996).

Utsumi A. Stylistic and Contextual Effects in Irony Processing. Proceedings of the 26<sup>th</sup> Annual Meeting of the Cognitive Science Society (CogSci2004), pp. 1369-1374 (2004).

Vealo T., Hao Y. Detecting Ironic Intent in Creative Comparisons. Proceedings of the 2010 conference on ECAI 2010: 19th European Conference on Artificial Intelligence, pp. 765-770 (2010).