Limitation des systèmes de détection computationnelle de l'ironie

L’ironie verbale est un phénomène difficile à définir. Dans l'avancement des recherches sur les algorithmes de détection des sentiments, ce type d'usage langagier pose plusieurs problèmes. En effet, il est difficile de conceptualiser un système en mesure d'identifier la présence d'ironie verbale dans un énoncé. Les systèmes développés actuellement démontrent peu de considération pour la nature linguistique du phénomène à étudier.

D’une part, le premier chapitre de ce mémoire abordera les recherches sur l'ironie qui exposent les différents enjeux sous-jacents à sa caractérisation. Pour Grice (1975), cet usage contrevient à la maxime conversationnelle de qualité stipulant qu'un locuteur doit toujours dire ce qu'il pense être vrai. Sperber & Wilson (1992), quant à eux, nuancent cette approche en mentionnant que cette définition néglige certaines formes d’ironie, comme les litotes et les hyperboles ironiques. Pour eux, ce type d'usage ferait plutôt écho à une connaissance partagée entre le locuteur et son interlocuteur. Ces définitions omettent certaines caractéristiques de l’ironie, notamment son aspect évaluatif. Concernant celui-ci, Alba-Juez & Attardo (2014) démontrent qu’un énoncé ironique peut communiquer des attitudes positives, négatives et, même, neutres.

D’autre part, dans le deuxième chapitre de ce mémoire, il sera question des méthodes computationnelles de détection de l’ironie. Ces dernières se basent généralement sur des caractéristiques superficielles de l’ironie verbale. Par exemple, pour Joshi et al. (2015), on pourrait détecter l’ironie en se basant sur l’incongruité implicite ou explicite d’un énoncé. Concrètement, si on retrouve la présence de valence négative et neutre dans une phrase, on peut soupçonner la présence d’ironie. D’autres algorithmes à base d’apprentissage machine, quant à eux, nécessitent de s’entraîner sur des corpus préalablement annotés. La mise en opposition des concepts présentés dans ces deux chapitres met en évidence des problèmes comme l’utilisation interchangeable de « sarcasme » et d’« ironie », une définition limitée à la valence négative communiquée par ce type de discours et l’élaboration d’un corpus basé sur une définition lacunaire. Finalement, dans le troisième chapitre de ce travail, je justifierai ces limitations et les raisons théoriques sous-jacentes à celles-ci. Pour ce faire, je testerai un modèle de langue, BERT (Devlin et al. 2018), ajusté finement sur des données d’entraînement tiré de la tâche de SemEval 2018 (Van Hee, 2018). En créant un corpus test tiré de la typologie d’énoncés ironiques proposée par Beals (1995), il me sera possible d’évaluer quels types d’ironie verbale sont plus difficiles à détecter par cet algorithme.

Cette mise en opposition des perspectives émergeants de la linguistique et de l’informatique finiront par mettre en évidence certaines caractéristiques relatives au traitement de l’ironie.