
Le style polémique dans les correspondances de Voltaire

Motasem Alrahabi^{*1}, Camille Koskas^{*2}, and Glenn Roe^{*3}

¹L'Observatoire des Textes, des Idées et des Corpus (ObTIC) – Sorbonne Université – France

²L'Observatoire des Textes, des Idées et des Corpus (ObTIC) – Sorbonne Université – France

³L'Observatoire des Textes, des Idées et des Corpus (ObTIC) – Sorbonne Université – France

Résumé

Résumé

La correspondance est un lieu d'échange et de dialogue qui offre un espace privilégié à l'expression des controverses, des polémiques et des débats [1]. Dans une perspective d'analyse automatique, nous nous proposons d'étudier les stratégies dialogiques inhérentes au style polémique dans la correspondance de Voltaire. Numérisé dans le cadre du projet Electronic Enlightenment à l'Université d'Oxford [2], le corpus de travail est composé d'environ 21000 lettres structurées en XML-TEI.

L'approche utilisée est symbolique et vise à annoter automatiquement les passages polémiques reflétant les postures et les positionnements des protagonistes autour de sujets débattus. En nous appuyant sur des travaux linguistiques autour de la subjectivité [3, 4, 5] et en traitement automatique des langues [6, 7, 8], nous avons mené une construction systématique de dictionnaires de patrons linguistiques à caractère polémique, organisés en quelques dizaines de classes : accords, désaccords, opinions, critiques, éloges, excuses, insultes, etc.

L'application finale permet de naviguer dans les lettres à travers les passages polémiques annotés. L'utilisateur a le moyen de filtrer les résultats par des mots clés et des métadonnées, et de les croiser avec des noms propres (personnes et lieux cités) automatiquement repérés à l'aide d'un modèle d'apprentissage automatique [9].

L'étude des résultats d'annotation nous a permis de mettre en avant la spécificité des marqueurs linguistiques observables qui accompagnent l'expression de la polémique dans le corpus. Dans cet espace de dialogue d'idées, nous avons pu constater chez Voltaire, à des degrés divers, une grande importance accordée aux éléments stylistiques de détournement du discours adverse : parodie, humour, sarcasme et satire. D'autres procédés annotés pourraient être caractéristiques du style polémique voltairien [10], notamment l'inscription de différentes voix opposées dans un espace polyphonique, et la manipulation savante de différentes formes du discours rapporté : le discours direct qui renforce l'aspect objectif du débat ; mais surtout le discours indirect où les propos de l'adversaire semblent d'abord tolérés pour vite être disqualifiés.

Les résultats des premières évaluations sont encourageants malgré certaines difficultés que nous soulignons : la nature très particulière de l'écriture épistolaire (orthographe ancienne,

*Intervenant

textes courts, changement thématique fréquent...) ; l'attribution des discours cités dans les lettres (distinguer les propos de l'écrivain de celui d'une personne citée [11]), les messages à caractère ironique ou allusif qui sont très difficiles à détecter automatiquement [12], etc.

Après une démonstration de l'outil sur le corpus de Voltaire, la présentation portera sur les caractéristiques du corpus (numérisation, structuration, modernisation...), sur la méthodologie de l'analyse automatique et son évaluation, et sur les perspectives.

Références:

Gavoille E., Guillaumont, F. (dir). *Conflits et polémiques dans l'épistolaire*, Presses universitaires François Rabelais, 2015.

Cronk, N., (2019), "Electronic Enlightenment: recreating the Republic of Letters", in Burrows, S. and Roe, G., eds., *Digitizing Enlightenment: Digital Humanities and the Transformation of Eighteenth century Studies*. Oxford University Studies in the Enlightenment: 55-72.

Kerbrat-Orecchioni C., (1980), " La polémique et ses définitions ", pp.3-40, in : Gelas N., Kerbrat-Orecchioni C., éd., *Le Discours polémique*, Lyon, Presses universitaires de Lyon.

Banfield, A., (1982), *Unspeakable Sentences*, Boston, Routledge and Kegan Paul.

Culioli, A., (1999), *Pour une linguistique de l'énonciation*, Tome 2, Paris, Ophrys.

Balahur, A., Mihalcea, R, Montoyo, A., (2014), "Computational approaches to subjectivity and sentiment analysis: Present and envisaged methods and applications", *Computer Speech & Language*, Volume 28, Issue 1.

Taboada, M., Brooke, J., Tofiloski, M., Voll, K., and Stede, M. (2011), "Lexicon-based methods for sentiment analysis". *Computational linguistics*, 37(2):267–307.

Alrahabi M., (2016), "E-Quotes : un outil de navigation textuelle guidée par les annotations sémantiques", in *TALN 2016*, Paris, France

Qi, Peng, Yuhao Zhang, Yuhui Zhang, Jason Bolton et Christopher D. Manning. 2020. "Stanza: A Python Natural Language Processing Toolkit for Many Human Languages". Dans *Proceedings of the 58th Annual Meeting of the Association for Computational Linguistics: System Demonstrations*, 101-108. Online : Association for Computational Linguistics. <https://doi.org/10.18653/v1/2020.acl-demos.14>.

Ibtissem Skander. *Le Dialogue polémique dans l'œuvre philosophique de Voltaire – étude stylistique* Thèse de doctorat, Université de Sfax (Tunisie), 28 octobre 2017.

Pareti, S., O'Keefe, T., Konstas, I., Curran, J.R., Koprinska, I., (2013), "Automatically Detecting and Attributing Indirect Quotations". In: *Empirical Methods in Natural Language Processing (EMNLP)*, Seattle

Hernández I., Rosso P., (2016), "Irony, Sarcasm, and Sentiment Analysis". Chapter 7 In: *Sentiment Analysis in Social Networks*, F.A. Pozzi, E. Fersini, E. Messina, and B. Liu (Eds.), Elsevier Science and Technology, pp. 113-128

