
Théâtre classique et difficultés numériques : quelques réflexions sur l'analyse dramatique assistée par ordinateur

Ioana Galleron*¹

¹LATTICE – Université Sorbonne Nouvelle - Paris 3, CNRS : UMR8094 – France

Résumé

Peut-on étudier le théâtre français de la période classique (1600-1800) avec des méthodes computationnelles ? La réponse à cette question est évidemment oui, mais l'objectif de cette communication est de souligner l'existence de plusieurs difficultés et problèmes qui rendent cette étude moins évidente qu'on ne peut le penser a priori. L'intervention tracera ainsi une sorte de " plan de travail " pour le développement des études numériques sur le genre et la période envisagée, tout en contribuant à la réflexion sur les changements épistémologiques impliqués par le développement de telles études.

1o La question des corpus

Parmi les différents genres littéraires des XVIIe et XVIIIe siècle, le théâtre connaît sans aucun doute la meilleure couverture en termes de numérisation. En effet, le ratio entre les documents préservés sur papier, et ceux qui ont été numérisés ne serait-ce qu'en mode image est largement favorable au théâtre, comme je le montrerai à partir d'une étude réalisée sur la base du catalogue de la BNF et de différents entrepôts numériques. L'écart est encore plus important lorsqu'on compare les ressources qui peuvent être manipulées par ordinateur (formats txt et xml), et celles qui ne peuvent être qu'affichées pour une lecture humaine (visionnage sous forme de pdf ou lecture exclusive à travers une plateforme en ligne).

Toutefois, il ne s'ensuit pas de là que le spécialiste du théâtre est privilégié. En effet, lorsqu'on essaie d'utiliser les ressources numériques " machine readable " on se confronte à d'autres questions, liées à leur correction, à leur interopérabilité et, non en dernier lieu, à leur représentativité. En d'autres mots, la constitution d'un corpus, fondé sur des choix explicites et systématiques, est loin d'être facile dans le domaine du théâtre français du XVIIe et XVIIIe siècle.

2o La question des observables

Comme le suggère la section précédente, une approche de type " distant reading " est sinon impossible, du moins difficile dans le cas des études sur le théâtre des siècles classiques. Une approche " smart data " plutôt que " big data " semble donc, à l'heure actuelle, plus adaptée pour les études théâtrales françaises, d'autant plus qu'elle s'inscrirait dans la mouvance actuelle de mise en place d'une " data-rich literary history " (Bode, 2018). En d'autres mots, le littéraire intéressé par l'analyse du théâtre doit envisager de créer ses propres observables (Landragin et Poudat, 2018), plutôt que d'utiliser des données pré-existantes.

*Intervenant

Or, les initiatives d'enrichissement des ressources restent actuellement limitées. Pour ne prendre qu'un seul exemple, l'identification des entités nommées reste très imparfaite avec les outils actuellement existants, au théâtre comme pour d'autres genres littéraires. Le progrès ne peut venir que de l'intérêt des littéraires pour de tels chantiers, qui demandent au préalable un travail d'annotation manuelle. Mais de tels travaux suscitent plusieurs questionnements, dont je donnerai plusieurs exemples à partir des pièces de théâtre.

3o La question des perspectives

Cette partie prolonge la précédente, en prenant appui sur l'observation que la définition des observables et l'automatisation de leur identification ont toujours pris appui, jusqu'à présent, sur des curiosités et des initiatives venues d'autres disciplines que les lettres, dans lesquelles l'étude du théâtre français des XVII^e et XVIII^e siècle est pourtant naturellement ancré. Pour le dire autrement, les littéraires se sont emparés de données et de méthodes élaborés par les linguistes, les informaticiens, parfois les géographes et les sociologues, mais se sont plus rarement attachés à l'opérationnalisation numérique de leurs propres objets d'études. On étudie les mots les plus fréquents, ou les mots-clés, parce que la linguistique a élaboré des protocoles pour ce faire, et que les littéraires ont pu tirer d'intéressantes conclusions pour leur discipline à partir de telles approches. On participe à l'effort de travail sur les entités nommées, parce qu'on a le sentiment qu'une nouvelle conceptualisation du personnage est possible dans le sillage de tels travaux. Mais il est plus rare que la démarche soit non pas celle d'un emprunt, mais plutôt celle d'une création – non pas dans le sens d'une création d'outil, ou d'interface, permettant d'étudier un aspect du texte littéraire, mais de modélisation d'un aspect du texte littéraire à étudier, menant à une réflexion sur la préparation éditoriale réaliser, puis à l'identification des outils appropriés pour mener cette étude. En d'autres mots, on assiste dans les études littéraires à la prévalence d'une pensée haptique (Rockwell, 2003) au détriment de l'émergence de "tinker toys" (W. McCarty, 2005). Sans dénier l'intérêt de la première (sur l'intérêt de laquelle v. Galleron et Idmhand, 2020), et sans opposer absolument les deux approches, il s'agit de proposer dans cette partie une réflexion sur la différence entre ce que les études théâtrales "traditionnelles" (pour les désigner avec une épithète commode, mais discutable) proposent comme programme d'étude du texte de théâtre, et ce que l'on peut actuellement réaliser grâce aux ressources numériques.

Bibliographie

Bode, Katherine. *A World of Fiction: Digital Collections and the Future of Literary History*. University of Michigan Press, 2018. doi:10.1353/book.59018.

Galleron, Ioana; Idmhand, Fatiha. "Why Go from Texts to Data, or The Digital Humanities as A Critique of the Humanities". *Word and Text. A journal of Literary Studies and Linguistics*, vol. X/ 2020, p. 53-69, http://jls.upg-ploiesti.ro/site_engleza/No_1_2020.html

Poudat Céline, Landragin Frédéric, *Explorer un corpus textuel. Méthodes - pratiques - outils*. De Boeck Supérieur, "Champs linguistiques", 2017, DOI : 10.3917/dbu.pouda.2017.01.

McCarty, Willard. *Humanities Computing*. Basingstoke: Palgrave MacMillan, 2005.

Rockwell, Geoffrey. "What Is Text Analysis, Really?". *Literary and Linguistic Computing* 18.2 (2003): 209-19. Available at <https://doi.org/10.1093/lc/18.2.209>. Accessed 16 September 2020.