

# Corneille et Molière

## Table ronde

Dans ses publications et son ouvrage : *Corneille dans l'ombre de Molière*, Dominique Labbé, spécialiste de statistique lexicale, affirme avoir contribué à résoudre une énigme scientifique en déterminant que certains chefs-d'œuvre de Molière (*Le Tartuffe*, *Dom Juan*, *Le Misanthrope*, *L'Avare*) sont en fait l'œuvre de Corneille. L'argument principal de cette découverte est l'étonnante similarité du vocabulaire des plus grandes pièces de Molière et de deux comédies de Corneille. Parce qu'elle a paru ébranler un mythe national, cette affirmation a provoqué un bouleversement qui a dépassé le seul monde des érudits pour atteindre la grande presse et les médias. Les conclusions de D. Labbé sont fondées sur la « distance intertextuelle », instrument statistique développé avec Cyril Labbé. Cette méthode mesure la différence globale de vocabulaire entre deux textes en déterminant leur différence relative en occurrences de mots : plus le nombre est bas, plus il est probable que les œuvres soient du même auteur. Les auteurs précités concluent que la distance lexicale de 16 pièces de Molière avec deux pièces anciennes de Corneille est suffisamment proche de zéro pour prouver que les textes sont de la même main. La polémique suscitée a mobilisé des statisticiens et des spécialistes du théâtre du XVII<sup>e</sup> siècle : les uns et les autres estiment que la statistique « ne devrait être utilisée que comme un auxiliaire complétant l'analyse littéraire et les données historiques », comme un « instrument exploratoire avec un objectif de description, mais en aucun cas comme une preuve ». Au cœur du débat se trouve donc la question fondamentale de l'usage qui est fait des instruments statistiques dans le domaine des lettres. Les historiens de la littérature, en outre, apportent dans la question de la paternité des œuvres de Molière des faits et des arguments d'ordre biographique, psychologique, historique, esthétique et stylistique.

### Dominique Labbé

CERAT-IEP - Université Pierre Mendès-France – Grenoble  
dominique.labbe@iep.upmf-grenoble.fr

Les travaux qui sont à l'origine de ce débat sont techniques, transparents et publiés. Nous n'avons pas une thèse, polémique ou pas, sur le sujet, mais une méthodologie dont un résumé de 20 lignes (demandé un peu tardivement) ne peut rendre compte. Nous pensons plus utile de donner quelque références.

Notre travail a été publié dans une revue internationale à comité de lecture :

Labbé Cyril, Labbé Dominique. « Inter-Textual Distance and Authorship Attribution Corneille and Molière ». *Journal of Quantitative Linguistics*. 8-3. December 2001. 213-231.

La distance intertextuelle a été présentée en français (avec une discussion de ses propriétés) :

Labbé Cyril, Labbé Dominique. « La distance intertextuelle ». *Corpus*. 2-2003. 95-118.

La méthode a été validée par plusieurs expériences techniquement irréprochables ; elle est utilisée par plusieurs chercheurs. Des publications sont en cours. Deux articles ont déjà paru :

Merriam Thomas. "Intertextual Distances between Shakespeare Plays, with Special Reference to Henry V (verse)". *Journal of Quantitative Linguistics*. 9-3. December 2002. 260-273.

Merriam Thomas. "An Application of Authorship Attribution by Intertextual Distance in English". *Corpus*. 2. 2003. 167-182.

Une synthèse concernant l'ensemble du dossier est accessible à :

<http://www.upmf-grenoble.fr/cerat/Recherche/PagesPerso/LabbeCorneilleMoliere.pdf>

(On trouvera sur cette page une série d'autres documents sur ce thème).

L'Université d'Oxford a mis en ligne toutes les pièces de Corneille, Molière et Racine — normalisées et lemmatisées par nos soins — avec d'autres corpus traités de la même manière, des publications et la plupart de nos logiciels : <http://ota.ahds.ac.uk/2466>.

Nous n'avons rien à vendre et nous n'utilisons aucune boîte noire.

### Jean-Marie Viprey

Archives, Textes, Sciences des Textes (EA 3187)

Université de Franche-Comté

Les statistiques lexicales sont-elles destinées à fournir des preuves ? Ne servent-elles pas plutôt à renforcer ou affaiblir des hypothèses, et surtout à les réorienter ? Une méthode probabiliste ne saurait, à elle seule, prouver qu'un auteur peut être tenu pour responsable de telle ou telle œuvre particulière.

Corneille et Molière sont intervenus dans un même champ littéraire, à des dates proches, et dans un genre très contraignant, la comédie (voire la comédie en vers). Corneille a mis la main à *Psyché*. Molière a écrit *Dom Garcie de Navarre* sous l'influence de la comédie héroïque cornélienne.

À mon sens, les résultats obtenus jusqu'ici accèdent plutôt la thèse communément admise : Molière invente peu à peu, et de façon décisive à partir de 1662 (*L'École des femmes*), une expression théâtrale radicalement nouvelle.

L'analyse lexicale massive, qui ne prend pas en compte les collocations ni *a fortiori* les configurations plus complexes, me paraît insuffisante pour caractériser une écriture. Par exemple, l'examen des interjections employées exclusivement dans les textes signés par Molière manifeste la référence spécifique de ce dernier à la *Commedia dell'Arte* (à l'exclusion de toutes les pièces de Corneille), ce qui affaiblit notablement la thèse présentée par Dominique Labbé.