

LES ÉTUDES CLASSIQUES

TOME LXXXII
2014

Langues anciennes et analyse statistique :
cinquante ans après
Distances textuelles et intertextualités

Édité par

Sabine FIALON, Dominique LONGRÉE

et Paul PIETQUIN

SOCIÉTÉ DES ÉTUDES CLASSIQUES
NAMUR

Avant-propos

Le 11 septembre 1961, le Conseil d'Administration de l'Université de Liège instituait le Laboratoire d'Analyse Statistique des langues anciennes. Le LASLA a donc fêté ses cinquante ans d'existence en 2011. À cette occasion, il a semblé utile et opportun de faire le point sur un des thèmes centraux des recherches menées par le LASLA depuis sa fondation, à savoir la question du calcul de la distance intertextuelle. Il s'agissait non seulement, dans une visée rétrospective, d'évaluer l'apport des méthodes développées depuis cinquante ans soit au LASLA, soit avec la collaboration du LASLA, mais aussi de lancer de nouvelles pistes pour les recherches à venir. Dans cette optique, le LASLA, en collaboration avec l'*Academia Belgica*, a organisé à Rome, les 19, 20 et 21 octobre 2011, des Journées d'études sur le thème « Langues anciennes et analyse statistique : cinquante ans après – distances textuelles et intertextualités ». La problématique de ces journées a été étendue à la question de l'intertextualité, dans la mesure où celle-ci peut parfois fonder des rapprochements ou des distances entre textes, comme c'est, par exemple, le cas dans le corpus hagiographique latin. Les travaux repris dans le présent recueil firent l'objet d'exposés et de discussions dans le cadre de ces journées.

Un point commun à toutes les études de ce recueil est une préoccupation méthodologique. Certains travaux partent des méthodes qu'ils testent sur divers textes, tandis que d'autres partent d'études de cas concrets pour voir comment les méthodes de l'analyse statistique des données textuelles peuvent apporter des réponses aux questions que ceux-ci posent. Mais ces diverses études visent toutes à faire progresser à la fois les méthodes d'analyse des textes et notre connaissance de ces derniers.

Un premier groupe d'articles a pour point commun de s'interroger sur la pertinence et l'efficacité des méthodes utilisées pour calculer les distances intertextuelles ou déceler des marques d'intertextualités entre œuvres littéraires. Dans un article qu'il dédie à Étienne Évrard, membre fondateur et inspirateur du LASLA, Étienne Brunet s'interroge sur les méthodes et outils à utiliser pour mesurer la connexion lexicale qu'il voit comme la forme positive de ce que représente en négatif le concept de

distance intertextuelle. Partant du cas concret de δόμος chez Euripide, confronté à *domus* chez Sénèque le Tragique, Marc Vandersmissen teste la méthode des cooccurrents spécifiques pour évaluer des distances entre deux textes rédigés dans deux langues différentes. S'inscrivant dans le cadre de recherches sur la longueur et la structure des phrases qui remontent aux origines même du LASLA, le travail de Dominique Longrée et Gérard Purnelle explore les possibilités offertes par le programme d'analyse syntaxique « LatSynt » pour développer un nouvel outil de calcul des distances intertextuelles à partir de la longueur des propositions subordonnées. S'appuyant sur une approche topologique des textes, Dominique Longrée et Sylvie Mellet montrent comment les variantes d'un motif peuvent constituer un outil à la fois pour caractériser un genre ou le style d'un prosateur latin, et pour calculer par là-même des distances entre genres ou spécificités d'auteurs. Marc Vandersmissen et Koen Vanhaegendoren présentent un logiciel développé spécifiquement à l'Université de Liège sur la base de fichiers de textes lemmatisés : celui-ci vise à détecter de nouvelles réminiscences potentielles de Xénophon dans l'œuvre d'Arrien, le premier étant reconnu comme le modèle admiré par le second. Edoardo d'Angelo offre enfin une réflexion sur les changements et progrès scientifiques de la métricologie médio-latine, ainsi que sur l'apport de la statistique et de l'informatique à ce domaine.

Un deuxième groupe d'articles propose différents exemples d'applications des méthodes de calcul de distance intertextuelle dans un domaine où les relations entre textes sont souvent particulièrement difficiles à établir, à savoir celui de l'hagiographie latine. Sabine Fialon montre ainsi comment les méthodes de l'analyse statistique des données textuelles peuvent apporter de nouveaux éléments de réflexion sur la notion d'*Africitas* appliquée à l'hagiographie africaine des II^e au VI^e siècles. En focalisant leur recherche sur une passion particulière, la *Passion de Sainte Marine*, Monique Goulet et Caroline Philippart de Foy s'interrogent sur ce que ces mêmes méthodes peuvent apporter à l'étude des distances entre des textes pré-carolingiens non normés, notamment pour distinguer les œuvres originales des traductions du grec. Jean Meyers démontre l'intérêt des méthodes d'analyse des données textuelles pour comparer dans le détail la langue et le style des cinq *Vies de Lambert* rédigées entre le VIII^e et le XII^e siècle. S'intéressant plus globalement à la littérature médio-latine, Innocenzo Mazzini montre comment une étude statistique, en l'occurrence celle de clausules métriques, pourrait remettre en question une fausse vérité, comme le fait que les auteurs médiévaux aient ignoré l'existence du texte de Lucrèce.

La thématique de la distance intertextuelle est donc ici abordée sous des angles divers, tant de par la variété des objectifs poursuivis (classification

de textes et mise en évidence des critères linguistiques ou stylistiques fondant des regroupements d'ordres thématique, générique ou chronologique ; recherches de modèles ou de réminiscences intertextuelles), que par les types de corpus étudiés (corpus homogènes, grecs ou latins, classiques ou médiévaux ; corpus hétérogènes, notamment chronologiquement) ou encore les outils utilisés (calculs statistiques de distance sur la base des fréquences, étude de cooccurrences, approches topologiques ...). Le présent recueil vise donc avant tout à promouvoir de nouvelles méthodes favorisant une meilleure compréhension non seulement des littératures grecque et latine classiques, mais également des œuvres médio- et néo-latines.

Dominique LONGRÉE