

28 février 2017

Dominique Labbé

Jean Racine, plume de l'ombre ?



Jean Racine (1639-1699)



Jean de la Chapelle (1651-1723)



Jean-Galbert Campistron (1656-1723)

Université de Neuchâtel

Séminaire Linguistique du français moderne: Linguistique de corpus

28 février 2017

Dominique Labbé

dominique.labbe@umrpacte.fr

(Laboratoire Pacte – CNRS – Université de Grenoble)

<http://www.pacte-grenoble.fr/blog/membres/labbe-dominique/>

Jean Racine, plume de l'ombre ?

Résumé

J. Racine a produit - sous son nom - onze tragédies dont certaines continuent à être jouées aujourd'hui. En 1677, il entre au service du roi et abandonne le théâtre ne sortant de sa réserve qu'à deux reprises (Esther et Athalie). Pourtant, en 1695, le Père Colonia - lui-même dramaturge - affirme que J. Racine continue à produire des pièces et que J.-G. Campistron est son prête-nom. L'attribution d'auteur assistée par ordinateur permet de vérifier la réalité de ces affirmations. Au total, 14 tragédies viennent s'ajouter au corpus racinien. Certaines ont été de très grands succès et méritent d'être redécouvertes.

Abstract

Between 1664 and 1677, J. Racine produced eleven tragedies under his name, some of which are still played today. In 1677, he entered the service of the king and abandoned the theatre. He only went back twice to present Esther and Athalie. However, in 1695 a Jesuit named Colonia, himself a dramatist, claimed that J. Racine was still ghost-writing some plays and that J.-G. Campistron was his nominee. A computer-assisted authorship attribution makes it possible to establish the veracity of these assertions. In total, 14 tragedies have thus been added to the Racinian corpus. Some of them have been very successful and deserve to be rediscovered.

Ce document est mis à votre disposition pour un usage non-commercial.

Toute mise en ligne ou reproduction, partielle ou totale, ne peut être faite sans l'autorisation de l'auteur.

“Je ne crois pas qu’un ‘dédoublément’ soit possible. Trop profondes sont les racines des œuvres, et leurs ramifications, lorsqu’elles paraissent variées, très différentes les unes des autres, ne sauraient résister à un véritable examen.”

Romain Gary. *Vie et mort d’Emile Ajar*.

Les organisateurs du Séminaire "Linguistique du français moderne" ont choisi cette année de traiter des corpus et ont accepté que, au cours de cette séance, nous abordions l’un des usages de ces collections électroniques de textes : la détermination de l’auteur d’une oeuvre anonyme ou d’origine douteuse. C’est une question classique en histoire littéraire (Love 2002¹) que l’informatique a profondément renouvelée comme nous allons le voir. Puisque le français moderne – sujet de ce séminaire - commence aux XVI-XVIIe, nous traiterons d’un cas de cette époque : l’activité de l’ombre de Jean Racine (1639-1699).

J. Racine est né orphelin et sans fortune. Il est mort riche et appartenant au premier cercle entourant le roi Louis XIV (Picard 1956 et Picard 1976). Cette carrière exceptionnelle se déroule en deux périodes.

De 1664 à 1677, il présente 10 pièces de théâtre (9 tragédies et une comédie²). Certaines furent de grands succès - *Andromaque* (1667), *Mithridate* (1672), *Iphigénie* (1674), *Phèdre* (1677) - et sont encore souvent jouées de nos jours.

En 1677, il se marie, devient historiographe du roi et abandonne le théâtre. Officiellement, il ne sort de sa réserve qu’à deux reprises à la demande de Mme de Maintenon, seconde épouse du roi Louis XIV (*Esther* 1689 et *Athalie* 1691).

Pourtant en 1695 – soit 18 ans après la retraite officielle de J. Racine -, un jésuite lyonnais, le Père Dominique Colonia, publie une pièce de théâtre intitulée *Juba*. Dans la préface, il écrit ceci :

L’incident qui fait le fonds de cette pièce, est peut-être la plus grande époque de l’histoire profane : c’est la ruine entière de la république romaine, et l’établissement de la monarchie universelle de Jules César. (...) **Cette aventure tragique a paru si propre pour le théâtre à l’illustre Monsieur Racine, qu’un de ses amis, qui depuis plusieurs années remplit si dignement sa place, lui a ouï dire qu’il était résolu de la traiter avant que de renoncer à la tragédie.**

On ne trouvera pas mauvais, j’en suis sûr, que nous prenions les devants, et que nous commençons par mettre les ombres du tableau, en attendant que ce grand maître veuille y mettre les couleurs (Colonia 1695).

En marge de l’exemplaire conservé à la bibliothèque de Lyon – accessible en ligne via Google books -, est mentionné manuscritement le nom de cet ami qui, depuis plusieurs années, remplit si dignement la place de J. Racine : Jean-Galbert Campistron (1656-1723)³. Les pièces présentées par J.G. Campistron sont présentées en annexe I (pour leur argument et leur réception : Basson et Labbé 2015).

¹ Les références bibliographiques sont placées à la fin de ce texte.

² Description du corpus en annexe 1 de cette conférence. Pour les normes de dépouillement : Labbé 2002.

³ Présentation dans Jones 1972, Grosperin 2004 ; Campistron (2002) ; Basson & Labbé 2015.

Cette affirmation de Colonia amène une série de questions :

- J. Racine travaille-t-il pour le théâtre en 1695, 18 ans après y avoir officiellement renoncé ?
- Que veut dire D. de Colonia quand il écrit que J.-G. Campistron "remplit si dignement [l]a place" de J. Racine ?

Dans les années 1950, les descendants de J.-G. Campistron ont remis, aux Archives départementales de la Haute-Garonne à Toulouse, les papiers de la famille contenant plusieurs documents appartenant à leur ancêtre, parmi lesquels figure *Juba* (Gérard 1955). Ce fonds, encore en cours de classement, contient aussi les manuscrits de quatre autres tragédies en alexandrins. Une a été jouée puis publiée sous le nom de J.-G. Campistron (*Virginie*). Trois autres sont inédites : les quatre premiers actes d'une tragédie (*Aétius*), représentée à la Comédie française en 1693, sous le nom de J.-G. Campistron; la totalité de *César*, tragédie en cinq actes jamais représentée ; enfin, des fragments du début d'une tragédie à l'état d'ébauche, *Tachmas*.

Ce *Juba* est-il celui que, d'après D. Colonia, J. Racine était en train de produire en 1695 ? Les autres pièces de J.-G. Campistron sont-elles de la même plume que celles de J. Racine ?

Jusqu'à maintenant, il était impossible de répondre à ces questions car il n'existait aucun moyen d'identifier l'auteur d'un texte anonyme ou d'origine douteuse (section 1). Il n'en est plus de même aujourd'hui grâce à de nouveaux outils informatiques qui seront présentés en section 2. Ces outils révèlent qu'une même plume a composé les tragédies présentées par J. Racine et par J.-G. Campistron mais aussi par un troisième "auteur" (section 3). Quelques objections ont été faites à ces attributions, nous y répondons (section 4) avant de conclure.

1. L'impossible identification des auteurs ?

L'histoire littéraire comporte de nombreuses énigmes comme celle posée par D. Colonia. La plus récente porte sur l'identité réelle de la romancière Elena Ferrante (Peras 2016). A chaque fois les critiques et les universitaires sont incapables de donner une réponse et se laissent facilement tromper. Il y a quelques années, Romain Gary en a offert une illustration intéressante.

Gary-Ajar

R. Gary (1914-1980) a été l'un des romanciers français les plus lus à travers le monde durant la seconde moitié du XXe siècle (Bellos 2010). Il a obtenu deux fois le prix Goncourt... qu'un écrivain ne peut recevoir qu'une fois ! La première fois en 1956 pour *les Racines du ciel* et la seconde fois en 1975 pour *la Vie devant soi*. Ce second roman a été publié sous le pseudonyme d'Emile Ajar avec la complicité d'un neveu (P. Pavlowitch) qui lui servit de prête-nom. Ajar a publié quatre romans qui ont tous été de gros succès de librairie, pendant que Gary continuait à publier de son côté. Pourtant, personne n'a éventé la supercherie.

En 1981, après la mort de R. Gary (1980), P. Pavlowitch a levé le voile, suivi d'une confession posthume de R. Gary parue la même année chez Gallimard (Gary 1981). Le monde littéraire en fut si mortifié qu'il a fait traverser à l'œuvre de Gary un purgatoire dont elle commence à sortir.

Des millions de personnes ont lu Gary *et* Ajar. Pratiquement aucune n'a décelé la même plume, même parmi les éditeurs, les critiques professionnels, les universitaires dont le métier consiste à analyser les textes. Dans son opuscule posthume, R. Gary cite les rares personnes qui ont relevé des parallèles étonnants entre les deux œuvres. Il s'agissait d'intuitions et toutes ces

personnes les ont abandonnées face aux dénégations de R. Gary. Une personne avait eu connaissance par accident de la supercherie :

"Une amie, venue chez moi à Majorque, avait vu sur mon bureau [le manuscrit de *Gros câlin*], avec le titre clairement marqué sur la couverture. Plus tard, lorsque le nom d'Emile Ajar prit le retentissement dont on trouvera la mesure en consultant les journaux de l'époque, c'est en vain qu'elle s'en alla partout, répétant que Romain Gary était l'auteur de l'œuvre, qu'elle avait vu, de ses yeux vu. On ne voulait rien savoir (...) Romain Gary était incapable d'avoir écrit cela" (p. 17).

Autrement dit, pour déterminer l'auteur réel d'un texte :

- on ne peut se fier aux "critiques de goût" (c'est génial ou c'est médiocre). En particulier, depuis plus de trois siècles, elles empêchent de prendre au sérieux le témoignage de Colonia ;

- les éditeurs, les critiques littéraires et les universitaires ne sont pas des sources fiables pour l'identification d'un auteur.

Or ce sont les seuls éléments dont on dispose pour la plus grande partie des pièces de théâtre du XVIIe, notamment celles présentées par J.-G. Campistron...

Considérons les années 1660-1700 (celles de J. Racine) : plus de la moitié des pièces de théâtre sont parues sous le nom d'intermédiaires et non pas sous ceux des écrivains qui les avaient composées (Labbé 2009a, chapitre VI). Le prête-nom était souvent un riche "comédien-poète". Dans cette même salle, en décembre 2009, le séminaire "mathématiques et société" a évoqué le cas de Molière (Labbé 2009b) Pour la période considérée, 25 comédiens – comme Molière - ont présenté plus de 300 pièces⁴. L'intermédiaire pouvait aussi être un riche financier. Par exemple, P. Quinault et J.-F. Regnard ont présenté sous leur nom des pièces écrites par Thomas Corneille (Labbé 2011).

A chaque fois, les troupes, les éditeurs, les gazettes ont présenté le prête-nom comme étant l'écrivain. Les "dix-septiémistes" ont emboîté le pas, faisant de ces intermédiaires des "auteurs". Pourtant, personne ne nie qu'il existe des "ressemblances" intrigantes entre certaines œuvres, et notamment entre celles présentées par J. Racine puis par J.-G. Campistron.

Racine- Campistron, un parallèle évident

Les contemporains de J.-G. Campistron ont été nombreux à faire le parallèle entre ses tragédies et celles de J. Racine. Ensuite, ce furent Voltaire puis V. Hugo. Chez les "dix-septiémistes", c'est quasiment un lieu commun :

"Il n'est pas besoin de s'arrêter longuement sur la manière dont Campistron s'efforce de donner une autre version de telle ou telle pièce de Racine, en inversant, transformant et renouvelant les cas tragiques." (Mazouer 1993)⁵.

Comme pour Ajar, ces parallèles n'ont jamais débouché à cause des mêmes obstacles. Comme pour Ajar, les troupes, les éditeurs, les journaux ont dit que J.-G. Campistron était l'auteur des pièces présentées sous son nom. De plus, les critiques et les universitaires unanimes ont affirmé que ces œuvres ne pouvaient sortir de la même main puisqu'ils jugeaient les unes géniales et les

⁴ Liste en annexe de Basson & Labbé 2015. Le recensement a été réalisé à l'aide de la base en ligne CESAR (Calendrier Electronique des Spectacles sous l'Ancien Régime et sous la révolution) et de : Mongrédien & Robert 1981.

⁵ Pour un aperçu de ces parallèles : Basson & Labbé 2015, p 17-19.

autres médiocres. Enfin, il n'existait jusqu'à aujourd'hui aucune méthode permettant de déterminer l'auteur d'un texte à l'aide des caractéristiques objectives de ce texte.

Aujourd'hui ce n'est plus le cas, l'informatique offre des outils semblables au microscope ou à la lunette astronomique qui permettent de voir l'infiniment petit ou l'infiniment lointain.

2. L'attribution d'auteur assistée par ordinateur

Grâce à l'informatique, l'attribution d'auteur assistée par ordinateur connaît un développement considérable (présentation dans Stamatatos, 2009 ; Jokers 2013 ; Koppel & Al., 2009 ; Savoy, 2012).

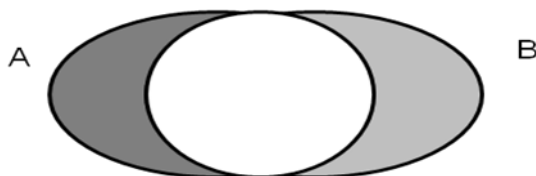
Notre méthode a été publiée pour la première fois en 2001. Un exposé en français, est consultable en ligne dans *Images des mathématiques*, revue des mathématiciens du CNRS destinée à un large public (Labbé & Labbé 2011). On y verra que l'attribution d'auteur n'est qu'une application particulière d'une procédure de classification automatique applicable aux grands corpus de textes, ce qui est l'objet du séminaire.

Voici les principaux points de cette méthode de classification et d'attribution d'auteur.

Méthode

Soit deux textes A et B. On superpose les deux textes et on compte le nombre de mots en commun (zone blanche) et le nombre de mots différents (zones grises de part et d'autre de la zone blanche). Puis les effectifs absolus sont convertis en nombres relatifs afin de donner un indice variant entre 0 et 1.

Le calcul de la distance intertextuelle



Une distance de 0,50 signifie qu'un mot sur deux est commun, l'autre étant différent ; une distance de 0,25 que les trois quarts des mots sont communs, etc.

Les expériences effectuées pour mettre au point la méthode ont permis d'identifier et de mesurer l'importance des principaux facteurs qui déterminent la distance entre textes (et leur classification automatique). Par importance décroissante, il s'agit de :

- une distance incompressible entre textes d'un même écrivain à la même époque, sur des thèmes voisins (pour le XVIIe siècle, elle varie entre 0.15 chez P. Corneille et 0.19 chez J. Racine) ;
- le genre : oral et écrit, prose, vers, comédie et tragédie, etc.
- l'écrivain,
- l'époque où a été rédigé le texte car chaque époque possède un vocabulaire particulier et le lexique de chaque écrivain évolue avec le temps,
- le thème (vocabulaire propre à ce thème, noms des personnages, lieux, principaux motifs).

On peut alors écrire :

$$D_{(A,B)} = f\{D_{\min(A,B)}, (Genre_A, Genre_B), (Auteur_A, Auteur_B), (Epoque_A, Epoque_B), (Thème_A, Thème_B)\}$$

Le symbole $f\{\}$ signifie que la distance est **fonction** des termes indiqués entre crochets. Ces termes sont nommés "variables", car ils peuvent être chiffrés. Pour donner un **poids** à chacune de ces variables, on cherche des cas où toutes les autres sont nulles ou négligeables (raisonnement "toutes choses égales par ailleurs").

Par exemple, en utilisant des textes appartenant au même genre (théâtre, poésie, roman, correspondance, etc.), écrits à la même époque, on peut isoler l'importance relative de l'écrivain et du thème. Parfois, on a la chance que deux écrivains contemporains traitent le même thème, en aveugles, dans le même genre. Cette situation est assez courante au XVII^e ; par exemple : les deux *Mère coquette* – T. Corneille (sous le nom de P. Quinault) et J. Donneau de Visé en 1665 - les deux *Bérénice* de P. Corneille et J. Racine (en 1670) ; les deux *Phèdre* (J. Racine et J. Pradon en 1677) ; les deux *Comte d'Essex* (T. Corneille et C. Boyer en 1678). Alors il ne reste plus que le facteur auteur.

De même, il est possible d'estimer le poids spécifique du temps sur la distance entre textes grâce à ceux écrits dans un même genre par les mêmes écrivains au cours de leur vie créatrice.

La principale conclusion est la suivante : dans un genre et à une époque donnée, la variable "auteur" l'emporte sur le thème et sur l'influence du temps, du moins quand celui-ci ne se compte pas en plusieurs dizaines d'années. Dès lors, **pour déterminer le véritable auteur d'un texte d'origine douteuse ou inconnue, il faut le confronter à d'autres – dont l'origine n'est pas douteuse – écrits dans un même genre et à la même époque.** Autrement dit : le théâtre doit être comparé au théâtre et les tragédies entre elles, les comédies entre elles, etc.

Par exemple, cette procédure désigne sans ambiguïté un auteur unique pour les romans parus sous les noms de R. Gary et E. Ajar (Labbé 2004).

Elle a été appliquée à *Aétius*, *Juba* et *Tachmas* comparées à 236 autres pièces de théâtre de la seconde moitié du XVII^e, par 32 "auteurs" différents (liste en annexe de Labbé 2014b). Elle révèle la même plume pour les tragédies présentées par J. Racine et J.-G. Campistron mais aussi celles d'un troisième "auteur" : Jean de la Chapelle (1651-1723, corpus présenté en annexe I).

3. J. Racine dans l'ombre de J. La Chapelle puis de J.-G. Campistron

Deux outils permettent d'identifier un seul écrivain pour les pièces parues sous ces trois noms.

Premier outil : une échelle standardisée des distances (présentation détaillée, notamment les méthodes d'étalonnage, dans Labbé 2014b). Dans le cas Racine-La Chapelle-Campistron (tableaux en annexe 2), les anomalies suivantes apparaissent par rapport à la situation attendue pour des écrivains contemporains, travaillant dans le même genre sur les mêmes thèmes :

- plusieurs distances proches de 0.20. En cas d'écrivains différents, ce phénomène ne peut se produire ;

- plusieurs distances inférieures ou égales à 0.23, ce qui a moins d'une chance sur mille de se produire en cas d'écrivains différents ;

- 86% des distances sont inférieures à 0.25 alors qu'on en attend moins de 5% en cas d'écrivains différents.

Ces constats conduisent au rejet des "plumes différentes" (même en collaboration) et à la conclusion d'un écrivain unique.

Deuxièmement, les classifications automatiques. L'annexe 3 présente le résultat de la méthode la plus simple (classification hiérarchique ascendante). A titre de "groupes témoins", on a introduit les tragédies de P. et T. Corneille contemporaines de celles de J. Racine.

L'algorithme procède à la construction d'une classe en regroupant les deux textes séparés par la distance la plus faible (ici deux pièces de P. Corneille : *Tite et Bérénice* et *Pulchérie*). Puis il recalcule les distances des autres textes par rapport à ce nouvel ensemble (par la moyenne arithmétique simple des distances), constitue une nouvelle classe, etc.

L'axe vertical du graphique indique à quelles distances les pièces ou groupe de pièces se rejoignent, ce qui donne une mesure de la plus ou moins grande homogénéité de chaque groupe.

A gauche du graphique, les tragédies contemporaines et P. et T. Corneille sont classées correctement et les deux groupes sont distincts. Il n'en est pas de même pour le groupe de droite.

Deux remarques à ce sujet.

Premièrement, la distance intertextuelle et la classification distinguent correctement les tragédies que les frères Corneille et J. Racine ont présentées à la même époque aux mêmes troupes et aux mêmes spectateurs (voir aussi : Labbé 2014). Etant donné que ces pièces sont en alexandrins et en cinq actes (sauf *Bérénice* qui n'en compte que trois), tous les facteurs autres que l'écrivain sont minimisés. A l'inverse, il s'écoule 26 ans entre *Andromaque* (1667) et *Aétius* (1693). Donc le facteur temps devrait contribuer à séparer J. Racine, J. La Chapelle et J.-G. Campistron...

Deuxièmement, Thomas est le frère cadet de Pierre, ils ont fait les mêmes études, ont épousé deux sœurs, vivent sous le même toit, élèvent ensemble leurs enfants et font bourse commune. L'histoire littéraire offre peu de cas d'une telle proximité entre deux écrivains. Même dans un pareil cas, notre méthode reconnaît sans ambiguïté leurs écrits respectifs.

En revanche, bien que s'étalant sur 26 ans, les tragédies présentées par J. Racine, J. La Chapelle et J.-G. Campistron sont classées ensemble, à des seuils qui indiquent sans ambiguïté un écrivain unique.

Autres indices

Beaucoup d'autres indices viennent renforcer cette conclusion : sens des mots les plus utilisés, combinaisons de mots, style. En ce qui concerne ce dernier point, la phrase est un aspect intéressant (Garette 1995 et Labbé & Labbé 2010).

Dans le théâtre versifié, la longueur de la phrase révèle la capacité d'un auteur à faire coïncider la période oratoire avec le vers (ou avec la césure de celui-ci). Le calcul compare la longueur de la phrase avec la longueur du vers (ici toujours l'alexandrin). Les deux figures en annexe IV rendent compte de ce calcul. Les phrases sont rangées par longueur croissante dans des classes d'intervalle égal à l'unité. Pour pouvoir comparer des corpus de longueurs différentes, les effectifs absolus des classes sont convertis en pourcentages. Le premier graphique compare les tragédies de J. Racine avec celles, contemporaines, de T. Corneille.

L'influence de la versification se voit clairement chez les deux, avec les trois mêmes modes principaux. Le mode de 9 mots correspond à un vers, celui de 18 à deux vers, etc. Malgré la contrainte de ce moule, J. Racine (trait gras) et T. Corneille (pointillés) sont bien distincts. Il y a pratiquement deux fois plus de phrases correspondant à un demi-vers ou à un vers unique chez J. Racine que chez T. Corneille. Chez le premier, les phrases courtes prédominent largement ; les

phrases moyennes et longues, voire très longues chez le second. Cela traduit des choix stylistiques différents.

Le deuxième graphique compare J. Racine et J.-G. Campistron. Les deux courbes se recouvrent quasi parfaitement. Cela confirme non seulement une seule plume mais un style remarquablement stable pendant plus d'un quart de siècle.

L'analyse peut aller plus loin. Le vocabulaire et les associations de mots caractéristiques de chaque sous-population de phrases permettent de dessiner un portrait stylistique de l'écrivain.

En définitive, tout converge pour désigner un écrivain unique pour les pièces parues sous les trois noms. La chronologie est cohérente : les trois "auteurs" se succèdent et produisent en moyenne un peu moins d'une tragédie par saison : J. Racine 1667-1677 (7 tragédies) ; J. de La Chapelle : 1681-82 (3 tragédies) ; J.-G. Campistron : 1683-1695 (8 tragédies plus trois inédites).

Cette attribution a soulevé quelques objections que nous allons examiner avant de conclure.

4. Quelques objections à l'attribution d'auteur par ordinateur

Parmi les objections, quelques-unes reviennent régulièrement, notamment à propos de l'erreur possible ou du caractère "réducteur" de la distance intertextuelle.

L'erreur est-elle possible ?

Il existe une marge d'erreur lorsqu'on travaille sur échantillons (cette marge décroît en raison inverse du carré de la dimension de l'échantillon). Nous ne travaillons pas sur échantillons mais sur des corpus exhaustifs dont chacun des mots est recensé à l'unité près. Il n'y a donc pas de "marge d'erreur" dans les calculs qui aboutissent à identifier un même auteur dans toutes ces pièces.

L'absence d'erreur vient aussi de ce que la méthode n'attribue un texte qu'à coup sûr. En contrepartie, elle ne répond pas toujours. Par exemple, *César*, la quatrième pièce des archives Campistron n'est pas attribuée. C'est ce qui a permis à la méthode de passer avec succès les nombreux tests en aveugle auxquels elle a été soumise depuis 19 ans (description en annexe de Labbé 2014a).

En mathématique appliquée, l'erreur peut enfin provenir de décisions erronées du chercheur. Le seul moyen de les éviter est d'éliminer toute intervention humaine dans le processus. Or les procédures présentées dans cette conférence sont automatiques. Notamment, il n'y a ni paramétrage ni apprentissage.

Tout chercheur appliquant les calculs sur les mêmes données aboutira aux mêmes résultats. Depuis notre première présentation sur le sujet en 1998, personne n'a contesté cette reproductibilité de la méthode qui est la meilleure garantie de sa qualité.

Convergence de tous les indices

Quant au caractère unique et réducteur de la mesure, la distance intertextuelle est la moins réductrice possible puisqu'elle considère tout le vocabulaire et donne à chaque mot un poids strictement équivalent au nombre de fois qu'il est employé.

De plus, c'est la convergence de tous les indices statistiques qui permet de conclure.

Enfin ces mesures convergentes sont étayées par plusieurs indices historiques que nous signalons dans notre ouvrage (Basson & Labbé 2015). Le principal d'entre eux est l'affirmation précise du père Colonia citée au début de cette conférence.

Le fonds Campistron des Archives départementales de Toulouse contient un autre indice important : une dizaine de lettres de la main de J.-G. Campistron. Dans notre ouvrage, nous reproduisons la première - contemporaine de *Virginie* (première tragédie présentée par Campistron) - et la dernière (postérieure à 1695). Toutes prouvent que Campistron maîtrise mal le français élémentaire. Il est impossible qu'il ait composé plus de 20 000 alexandrins.

Naturellement, cela ne l'a pas empêché de faire une belle carrière et de "tenir dignement la place de J. Racine", selon l'expression de D. de Colonia.

Pourquoi ce système du prête-nom ?

Autre objection : puisque ces pièces plaisaient à la cour et à la ville, il n'y avait aucune raison de se cacher derrière un prête-nom.

Revenons à notre propos de départ : J. Racine cesse officiellement de produire pour le théâtre quand il entre dans la maison du roi en 1677.

J. de La Chapelle et J.-G. Campistron suivent un chemin identique. En 1682, le premier entre au service du prince de Conti et abandonne le théâtre. J.-G. Campistron est secrétaire général aux galères, de 1692 (ou 1693) à 1709, et cesse toute production théâtrale durant cette période.

Dans le même sens et à la même époque, on trouve les cas d'Edmée Boursault (receveur des impôts) et de Gaspard Abeille (secrétaire du maréchal de Luxembourg) (Voir leurs portraits dans Basson & Labbé 2015).

Il y a donc bien une obligation de réserve qui pèse sur ce qu'on pourrait appeler les "hauts fonctionnaires" de l'époque.

Cependant, une pièce à succès peut rapporter gros, même en partageant le revenu avec un prête-nom. Tant que les deux associés restent discrets, l'écrivain peut tenir son rang et se ménager des revenus annexes non-négligeables. C'est ce que font E. Boursault (prête-nom : le comédien R. Poisson), G. Abeille (prête-nom le comédien La Thuilerie) ou J. Racine avec J. La Chapelle puis J.-G. Campistron.

En fait, le système du prête-nom était quasi général comme le suggère l'existence des nombreux comédiens poètes signalée au début de cette conférence et notamment les associations de P. Corneille avec Molière ou de son frère Thomas avec Montfleury puis Hauteroche.

Ce système s'explique d'abord par des raisons structurelles tenant à l'absence de propriété intellectuelle et commerciale ou de droit des sociétés. Les troupes n'ayant pas la personnalité juridique, elles ne peuvent passer contrat avec les écrivains pour se garantir l'exclusivité des pièces. Pour cela, un acteur ou un intermédiaire de confiance négocie avec l'écrivain et assume l'exploitation de la pièce. Cet achat (ou avance) permet aussi de se constituer une réserve de textes, spécialement les comédies qui plaisent au public populaire et comblent des vides durant la saison.

Pourquoi ces pièces sont-elles oubliées ou méprisées ?

Enfin dernière objection : les pièces de J. Racine seraient géniales, celles de J. de La Chapelle ou de J.-G Campistron seraient médiocres, donc ce ne peut pas être le même auteur... Nous avons déjà dit ce qu'il faut penser de ces "critiques de goût" à propos du cas Gary-Ajar.

A leur parution, les pièces J. de La Chapelle ou de J.-G Campistron ont été des succès comparables à ceux des pièces présentées par J. Racine avant 1677. Certaines sont entrées dans le répertoire : *Cléopâtre* (La Chapelle) ou *Andronic* (Campistron) ont été reprises régulièrement jusqu'à la Révolution.

Et puis elles ont sombré dans un oubli qui s'explique de plusieurs façons.

L'oubli s'explique d'abord par rôle des "autorités" en matière littéraire (critiques, grands auteurs, universitaires). En effet, non seulement, le lecteur n'a pas les moyens d'identifier l'auteur d'un texte mais il lui est difficile de "juger" la valeur de celui-ci. En général, il fait confiance aux critiques de goût formulées par les "autorités". Or, depuis plus de trois siècles celles-ci - Boileau, Voltaire, Hugo et, après eux, les "dix-septémistes" - disent : Racine est un génie ; Campistron (ou La Chapelle), des médiocres. Mais ces jugements reposent sur une erreur concernant l'identité de l'auteur !

La seconde raison tenait au fait que la plupart de ces pièces étaient introuvables n'ayant pas été rééditées depuis plus de deux siècles. Mais aujourd'hui, les éditions originales sont numérisées et accessibles en ligne grâce à Google books et Gallica. De plus, grâce à Paul Fièvre, la plupart ont été transcrites en français contemporain et sont disponibles gratuitement sur le site : <http://www.theatre-classique.fr/>

Pendant, la difficulté n'est qu'à demi résolue. Comme Georges Couton le remarquait à propos des dernières pièces de Corneille :

La lecture n'assure à un théâtre qu'une vie diminuée et précaire. Des pièces non jouées acquièrent inévitablement la réputation de pièces médiocres, qui ne méritent que d'être lues, puis qui ne méritent pas même d'être lues. (...) L'incuriosité du théâtre amène l'incuriosité de la lecture. (Couton présentation de Corneille 1980, p. LXXXI-LXXXII. Voir aussi Kerr 1994).

Ainsi le conformisme a fait son œuvre. Et l'on pense souvent que les contemporains se sont trompés en faisant un triomphe aux pièces de La Chapelle ou de Campistron.

Ne serait-il pas temps de relire ces textes sans ces préjugés qui viennent, avant tout, d'une erreur sur l'identité de celui qui les écrites ?

Conclusion

Au XVIIe siècle, l'"auteur" d'une pièce de théâtre est la personne qui a négocié cette pièce avec les comédiens, touché les droits puis assuré la publication en cas de succès. Son nom figure sur les affiches, le registre de la troupe, la couverture du livre. Dans plus de la moitié des cas, cet "auteur" était un intermédiaire, prête-nom de l'écrivain qui avait écrit la pièce. Il demeure donc un grand nombre d'œuvres "orphelines" ou douteuses sur lesquelles les recherches se poursuivent.

Contrairement à ce qu'avait affirmé R. Barthes en 1968, l'écrivain n'est pas mort. Il est bien présent dans ses œuvres, les moyens existent pour l'identifier et surtout pour en dresser le portrait lexical, stylistique, thématique. Nous l'avons suggéré brièvement à propos des phrases.

Les nouveaux outils d'attribution d'auteur seront précieux pour les historiens, notamment pour ceux qui revisiteront le théâtre du XVII^e sans s'arrêter aux critiques de goût qui encombrant les études littéraires. Au-delà de cette question, les mathématiques appliquées peuvent être utiles à toutes les sciences humaines et en premier lieu à la linguistique de corpus. Nous en avons évoqué quelques aspects, ici même lors du séminaire *Mathématiques et société* de novembre 2010 (Labbé 2010).

Dans le cas d'espèce, tous les indices statistiques et les données historiques convergent vers la même conclusion : une seule plume a écrit les 21 tragédies présentées, entre 1667 et 1693 successivement sous les noms de : J. Racine, J. de La Chapelle et J.-G. Campistron.

Le "corpus racinien" fait plus que doubler de volume. De plus, on dispose maintenant de manuscrits à différents stades d'élaboration, ce qui devrait permettre de mieux comprendre le processus de création de toutes ces pièces.

Naturellement, les tests statistiques ne reconnaissent pas l'écrivain, ils indiquent simplement que les textes présentés par J. Racine (entre 1667 et 1677), par J. de La Chapelle et par J.-G. Campistron ont été écrits par la même main. L'identification de cet écrivain unique passe par l'examen d'autres indices, notamment historiques. Par exemple, étant donné leurs dates de naissance, J.-G. Campistron (1656-1723) et J. de La Chapelle (1651-1723) ne peuvent être cet écrivain puisqu'ils étaient trop jeunes au moment où parurent les pièces présentées par J. Racine et qu'ils n'étaient pas présents à Paris. Si J. Racine a écrit les pièces parues sous son nom, alors il a aussi composé les 12 tragédies présentées par J. de La Chapelle puis par J.-G. Campistron, ainsi que *Juba* et *Tachmas*.

Une question demeure concernant les oeuvres présentées sous le nom de J. Racine.

Il y a un noyau central composé des 7 tragédies créées entre 1667 (*Andromaque*) à 1677 (*Phèdre*) – qui sont aussi les plus grands succès – auquel s'ajoutent donc les trois tragédies de J. de La Chapelle et les onze de J.-G. Campistron. A l'écart de ce bloc, on trouve d'une part les deux premières tragédies présentées par J. Racine - *la Thébàide* (1664) et *Alexandre* (1665) - et, d'autre part, les deux dernières (*Esther*, 1689 et *Athalie*, 1691). Racine était-il un "caméléon" pour reprendre la formule d'Alain Viala (1990) ? Ou bien son œuvre est-elle le fruit de plusieurs collaborations ? La question reste en suspens pour l'instant car nos outils ne peuvent y répondre à coup sûr, et que nous ne nous écarterons pas de cette règle que nous avons respectée depuis 20 ans.

Remerciements.

Les animateurs du séminaire "Linguistique du français moderne", Jacques Savoy et l'Université de Neuchâtel ont organisé cette conférence. Les archives de la Haute-Garonne à Toulouse ont permis à J.-C. Basson (Université de Toulouse-Le Mirail) de cliquer les manuscrits du fonds Maniban-Campistron. Jean-Charles Basson a également collaboré à la transcription de ces textes et à leur présentation. Cyril Labbé (Laboratoire d'Informatique de Grenoble) a aidé à la mise au point des logiciels. Edward Arnold, Guy Bensimon, Jean-Guy Bergeron, Mathieu Brugidou, Pierre Hubert, Nelly & Jean Leselbaum, Xuan Luong, Thomas Merriam, Denis Monière, Jacques Picard, André Pibarot et Mathieu Ruhlman ont collaboré à la mise au point de la méthode d'attribution d'auteur.

Bibliographie

NB : nos travaux sont en ligne dans les Archives ouvertes du CNRS et ResearchGate (les liens sont accessibles à partir de notre page personnelle).

- Barthes Roland (1968). La mort de l'auteur. *Œuvres complètes*. II. Paris : Seuil, 1994, p 491-495.
- Basson Jean-Charles & Labbé Dominique (2015). *Jean Racine. Aétius, Juba, Tachmas. Tragédies inédites transcrites et présentées par Jean-Charles Basson et Dominique Labbé*. Montréal : Monière-Wollank Editeurs (Consultable en ligne sur ResearchGate).
- Bellos David (2010). *Romain Gary : A Tall Story*. London : Harvil Secker.
- Bertaud Madeleine et Niderst Alain (dir.). *Onze études sur la vieillesse de Corneille dédiées à la mémoire de Georges Couton*. Paris : Klincksieck, 1994.
- Campistron Jean-Galbert (2002). *Tragédies (1684-1685). Arminius Andronic Alcibiade*. Edition établie, présentée et annotée par Jean-Philippe Groperrin et Jean-Noël Pascal. Toulouse : Société de Littératures Classiques.
- Colonia Dominique de.(1695). *Juba*. Lyon : Guerrier.
- Corneille (1980). *Œuvres complètes*. Paris : Gallimard.
- Garette Robert (1995). *La phrase de Racine. Etude stylistique et stylométrique*. Toulouse : Presses universitaires du Mirail.
- Gary Romain (1981). *Vie et mort de Emile Ajar*. Paris : Gallimard.
- Gérard Pierre (1955). Une tragédie inédite de Campistron. Juba, roy de Mauritanie. *Actes du 11e congrès d'études de la fédération des sociétés académiques et savantes Languedoc-Pyrénées - Gascogne*. Albi : Imprimerie des orphelins apprentis, p. 36-42.
- Groperrin Jean-Philippe (Ed 2004). *Campistron et consorts : tragédie et opéra en France (1680-1733)*. Littératures classiques. 52, automne 2004.
- Jockers Matthew L. (2013). *Macroanalysis: Digital Methods and Literary History*. Urbana: Univ. of Illinois Press, 2013.
- Jones Dorothy F (1972). *Jean de Campistron: a study of his life and work*. University Mississippi Romance monographs.
- Kerr Cynthia (1994). Opération sauvetage : la mise en scène des dernières pièces de Corneille. In Bertaud Madeleine et Niderst Alain (dir.). *Onze études sur la vieillesse de Corneille dédiées à la mémoire de Georges Couton*. Paris : Klincksieck, 1994, p. 171-185.
- Koppel Moshe, Schler Jonathan & Argamon Shlomo (2009). Computational Methods in Authorship Attribution. *Journal of the American Society for Information Science and Technology*. 2009, 60-1, p. 9-26.
- Gary Romain (1981). *Vie et mort de Emile Ajar*. Paris : Gallimard.
- Labbé Cyril & Labbé Dominique (2001). Inter-Textual Distance and Authorship Attribution Corneille and Molière. *Journal of Quantitative Linguistics*. 8-3, december 2001, p. 213-231.
- Labbé Cyril & Labbé Dominique (2010). Ce que disent leurs phrases. In Bolasco Sergio, Chiari Isabella, Giuliano Luca. *International Conference Statistical Analysis of Textual Data*. Rome : Edizioni Universitarie di Lettere Economia Diritto, 1, pp.297-307.
- Labbé Cyril & Labbé Dominique (2011). La classification des textes. Comment trouver le meilleur classement possible au sein d'une collection de textes ? *Images des mathématiques*. La

- recherche mathématique en mots et en images. (<http://images.math.cnrs.fr/La-classification-des-textes.html>). 28 mars 2011.
- Labbé Dominique (2002). La lemmatisation des grandes bases de textes. Un exemple : Corneille, Molière et Racine. *Colloque L'édition électronique en littérature et dictionnaire, évaluation et bilan*. Rouen : 17-21 juin 2002.
- Labbé Dominique (2004). *Romain Gary et Emile Ajar*. Grenoble : Cerat-IEP, mai 2004.
- Labbé Dominique (2009a). *Qui a écrit Tartuffe ?* Montréal : Monière et Wollank. Réédition : *Si deux et deux sont quatre Molière n'a pas écrit Don Juan*. Paris : Max Milo.
- Labbé Dominique (2009b). *Qui a écrit Dom Juan ? Molière est-il l'auteur des pièces parues sous son nom ?* Séminaire Mathématiques et société. Université de Neuchâtel : 9 décembre 2009.
- Labbé Dominique (2010). *Le calcul du sens des mots. La lexicologie assistée par ordinateur*. Séminaire Mathématiques et société. Neuchâtel, 3 novembre 2010.
- Labbé Dominique (2011). *Comédiens et écrivains au XVIIe siècle. A la redécouverte des frères Corneille*. Conférence invitée. Séminaire de stylistique française. Université de Cologne. Jeudi 9 juin 2011.
- Labbé Dominique (2014a). Identification de l'auteur d'un texte (Hugo, Lamartine, Musset et Vigny). Séminaire : *L'œuvre et son auteur : problèmes d'attribution*, 21 mai 2014, Université de Lille.
- Labbé Dominique (2014b). *Qui a écrit Aëtius, Juba et Tachmas ? Une attribution d'auteur par ordinateur*. Grenoble : Pacte.
- Love Harold (2002). *Attributing Authorship: an Introduction*. Cambridge. Cambridge University Press.
- Mazouer Charles (1993). La mémoire des classiques chez Regnard et Campistron. *Les lieux de mémoire et la fabrique de l'œuvre. Biblio 17 (Papers on French Seventeenth Century Literature)*, 80, p. 277-293.
- Mongrédien Georges et Robert Jean (1981). *Les comédiens français du XVIIe siècles. Dictionnaire biographique suivi d'un inventaire des troupes (1590-1710)*. Paris : Eds du CNRS.
- Monière Denis & Labbé Dominique (2006). "L'influence des plumes de l'ombre sur les discours des politiciens". In Condé Claude et Viprey Jean-Marie. *Actes des 8e Journées internationales d'Analyse des données textuelles*. Besançon, II, p. 687-696.
- Pavlowitch Paul (1981). *L'homme que l'on croyait*. Paris : Fayard.
- Peras Delphine (2016). Elena Ferrante, énigme littéraire fascinante. *L'Express*, 26 juin 2016.
- Savoy Jacques (2012). Authorship Attribution: A Comparative Study of Three Text Corpora and Three Languages. *Journal of Quantitative Linguistics*. 19(2), 2012, p. 132-161.
- Stamatatos Efsthios (2009). A survey of modern authorship attribution methods. *Journal of the American Society for information Science and Technology*. 60-3, p. 538-556.
- Truchet Jacques (1992). Tiridate. *Théâtre du XVIIe siècle*. Gallimard (La pléiade), 1992, tome III, p. et p. 1279-1284.
- Viala Alain (1990). *Racine, la stratégie du caméléon*. Paris : Seghers.

ANNEXE I. Les corpus (avec les longueurs en mots)

1. Jean Racine

Titre	Genre	Création	Longueur	Vocabulaire
La Thébaine	Tragédie	20 juin 1664	13 813	2 256
Alexandre le Grand	Tragédie	4 décembre 1665	13 864	2 383
Andromaque	Tragédie	17 novembre 1667	15 076	2 441
Britannicus	Tragédie	13 décembre 1669	15 387	2 748
Bérénice	Tragédie	21 novembre 1670	13 242	2 254
Bajazet	Tragédie	1er janvier 1672	15 297	2 598
Mithridate	Tragédie	23 décembre 1672	15 091	2 650
Iphigénie	Tragédie	18 août 1674	15 782	2 753
Phèdre	Tragédie	1er janvier 1677	14 394	2 918
Esther	Tragédie	26 janvier 1689	11 147	2 518
Athalie	Tragédie	5 janvier 1691	15 492	2 989
			166 626	10 145

Source : Paul Mesnard. *Oeuvres de Jean Racine*. Paris : Hachette, 1885 (Les Grands écrivains de la France)

2. Jean de la Chapelle

Titre	Genre	Création	Longueur	Vocabulaire
Zaïde	Tragédie	26 janvier 1681	13 048	2 314
Cléopâtre	Tragédie	12 décembre 1681	12 158	1 342
Téléphonte	Tragédie	26 décembre 1682	12 425	1 342
			37 631	2 138

Transcrites en français contemporain par D. Labbé d'après : *Cléopâtre* : La Haye : A. Moetjens, 1683. *Téléphonte* : Paris : Thomas Guillain, 1683. *Zaïde* : Paris : Jean Ribou, 1681.

3. Jean-Galbert Campistron

Titre	Genre	Création	Longueur	Vocabulaire
Virginie	Tragédie	12 février 1683	13 430	1 349
Arminius	Tragédie	19 février 1684	13 478	1 455
Andronic	Tragédie	8 février 1685	12 457	1 400
Alcibiade	Tragédie	28 décembre 1685	13 182	1 389
Acis et Galatée	Pastorale	septembre 1686	4 369	799
Achille et Polixène	Tragédie musicale	7 novembre 1688	5 990	924
Phocion	Tragédie	16 décembre 1688	10 713	1 315
Adrien	Tragédie	11 janvier 1690	12 835	1 499
Tiridate	Tragédie	12 décembre 1691	11 488	1 330
Aétius	Tragédie	28 janvier 1693	10 368	1 290
Alcide	Opéra	3 février 1693	5 563	926
Pompeia	Tragédie	1693	11 881	1 469
Juba	Tragédie inédite	1695	9 757	1 367
César	Tragédie inédite	-	15 322	1 716
Tachmas	Tragédie inédite	-	4 039	756
Total			183 085	4 233

Transcrites d'après les éditions originales des œuvres, les éditions de J. Truchet (*Tiridate*) et de J.-P. Gossperin (*Arminius*, *Andronic* et *Alcibiade*).

ANNEXE II.

1. Distances entre les pièces présentées entre 1683 et 1693 par J.-G. Campistron et celles de J. Racine entre 1667 et 1677.

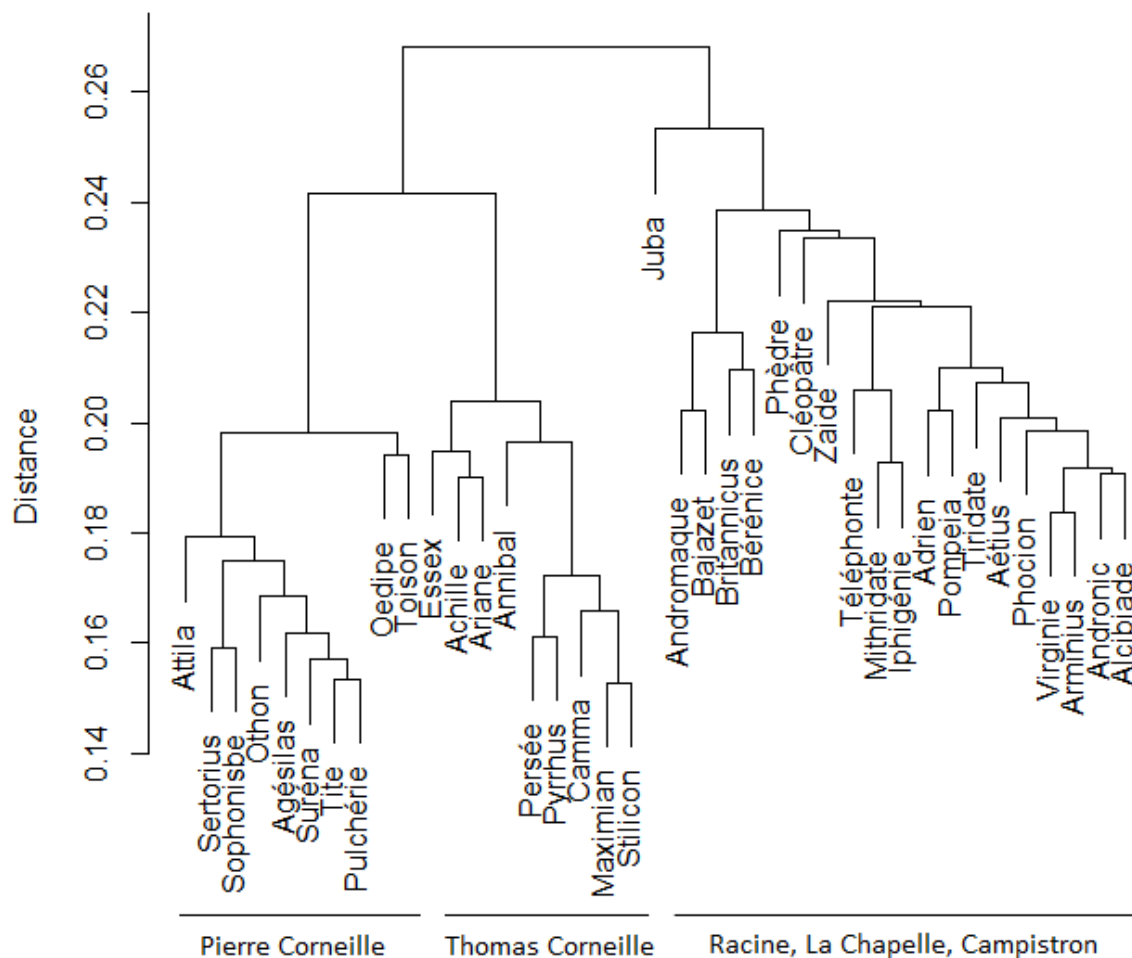
Racine Campistron	Andromaque (1667)	Britannicus (1669)	Bérénice (1670)	Bajazet (1672)	Mithridate (1672)	Iphigénie (1674)	Phèdre (1677)	Moyenne
Virginie (1683)	0,237	0,236	0,240	0,234	0,209	0,212	0,235	0,229
Arminius (1684)	0,237	0,232	0,228	0,230	0,201	0,211	0,232	0,225
Andronic (1685)	0,233	0,229	0,226	0,227	0,203	0,223	0,236	0,225
Alcibiade (1685)	0,241	0,233	0,236	0,233	0,209	0,216	0,230	0,228
Phocion (1688)	0,249	0,246	0,257	0,238	0,223	0,232	0,227	0,239
Adrien (1690)	0,265	0,270	0,271	0,265	0,244	0,237	0,238	0,256
Tiridate (1691)	0,242	0,234	0,231	0,248	0,224	0,234	0,236	0,236
Pompéia (1692)	0,251	0,251	0,249	0,242	0,216	0,234	0,237	0,240
Aétius (1693)	0,265	0,240	0,233	0,248	0,228	0,241	0,255	0,244
Juba (1695)	0,290	0,271	0,284	0,258	0,242	0,259	0,277	0,269
Moyenne	0,251	0,244	0,246	0,242	0,220	0,230	0,240	0,239

2. Distances entre les trois tragédies présentées par J. de La Chapelle et celles présentées par J. Racine et J.-G. Campistron.

La Chapelle	Zaïde (1681)	Téléphonte (1683)	Cléopâtre (1683)	Moyenne
J. Racine				
Andromaque (1667)	0,241	0,240	0,246	0,242
Britannicus (1669)	0,230	0,239	0,236	0,235
Bérénice (1670)	0,240	0,234	0,249	0,241
Bajazet (1672)	0,235	0,242	0,238	0,238
Mithridate (1672)	0,225	0,204	0,221	0,217
Iphigénie (1674)	0,234	0,207	0,231	0,224
Phèdre (1677)	0,238	0,238	0,246	0,240
Moyenne Racine	0,235	0,229	0,238	0,234
J.-G. Campistron				
Virginie (1683)	0,212	0,211	0,228	0,217
Arminius (1684)	0,213	0,211	0,221	0,215
Andronic (1685)	0,218	0,194	0,236	0,218
Alcibiade (1685)	0,205	0,221	0,223	0,216
Phocion (1688)	0,230	0,226	0,246	0,234
Adrien (1690)	0,242	0,233	0,248	0,241
Tiridate (1691)	0,228	0,220	0,251	0,233
Pompéia (1692)	0,226	0,222	0,244	0,230
Aétius (1693)	0,218	0,228	0,237	0,228
Juba (1695)	0,258	0,267	0,259	0,261
Tachmas (?)	0,219	0,227	0,249	0,232
Moyenne Campistron	0,225	0,223	0,259	0,229
Moyenne générale	0,229	0,225	0,239	0,231

ANNEXE III

Dendrogramme de la classification hiérarchique ascendante sur le corpus des pièces présentées par J. Racine, P. et T. Corneille, J. de La Chapelle et J.-G. Campistron (méthode de la moyenne).



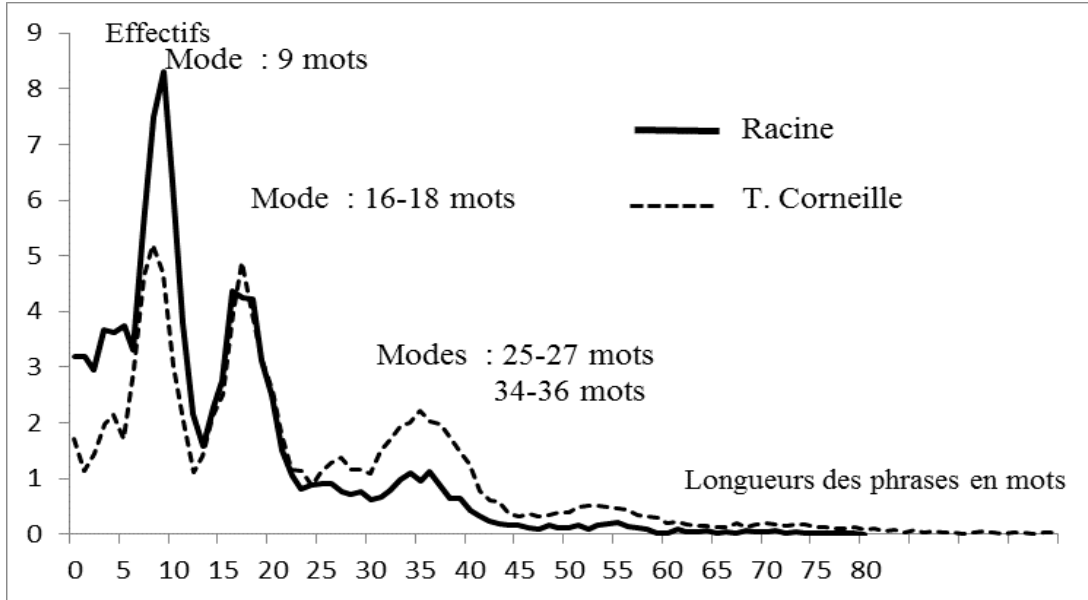
Pierre Corneille : *Œdipe* (1659), *la Toison d'or* (1661), *Sertorius* (1662), *Sophonisbe* (1663), *Othon* (1664), *Agésilas* (1666), *Attila* (1667), *Tite et Bérénice* (1670), *Pulchérie* (1672), *Suréna* (1674).

Thomas Corneille : *Stilicon* (1660), *Camma* (1661), *Persée et Démétrius* (1661), *Maximian* (1662), *Pyrrhus* (1663), *la Mort d'Annibal* (1669), *Ariane* (1672), *la Mort d'Achille* (1673), *le Comte d'Essex* (1678)

Racine, La Chapelle & Campistron, voir Tableaux en annexe ci-dessus.

ANNEXE IV

1. Histogramme des longueurs de phrases dans les tragédies de J. Racine et T. Corneille (effectifs en pourcentage du total des phrases).



2. Histogramme des longueurs de phrases dans les tragédies de J. Racine et de J.-G. Campistron

