

TPE : L'AFFAIRE CORNEILLE-MOLIERE

Une étude statistique peut-elle permettre, ou non, d'attribuer une œuvre à un auteur ?

Disciplines : Mathématiques et Littérature

Par Victor Eiferman, Pierre-Louis Jeauffroy et Guillaume Soupre

Lycée Carnot, année 2009-2010

SOMMAIRE

<u>INTRODUCTION</u>	p.4
<u>PARTIE 1 : CORNEILLE ET MOLIERE</u>	p.5
1) PIERRE CORNEILLE ET JEAN-BAPTISTE POQUELIN	p.5
2) LES DEBUTS DE L'AFFAIRE ET SON EVOLUTION	p.6
3) ARGUMENTS BIOGRAPHIQUES ET STYLISTIQUES	p.7
4) CONCLUSIONS	p.10
<u>PARTIE 2 : L'AFFAIRE ET LES MATHEMATIQUES</u>	p.11
1) LA STYLOMETRIE	p.11
2) L'ETUDE DE DOMINIQUE LABBE	p.11
a) Le calcul de la distance intertextuelle	p.12
b) Les combinaisons de mots	p.16
c) Le sens des mots	p.17
3) CONCLUSIONS	p.17
<u>PARTIE 3 : ROMAIN GARY – EMILE AJAR</u>	p.19
1) L'AFFAIRE GARY-AJAR	p.19
2) <i>LA VIE DEVANT SOI</i> ET <i>CLAIR DE FEMME</i>	p.20
a) <i>La Vie devant soi</i>, Momo et Rosa	p.20
b) <i>Clair de femme</i>, la rencontre	p.23
c) Conclusion	p.25

3) L'EXPERIENCE D'ATTRIBUTION D'AUTEUR	p.25
<u>CONCLUSION</u>	p.30
<u>LEXIQUE</u> (les mots suivis d'une astérisque y sont définis)	p.31
<u>ANNEXES</u>	
1) E-MAIL DE DENIS BOISSIER (cf. p.6)	p.32
2) ELOGE DE CORNEILLE PAR RACINE (cf. p.5)	p.34
3) ENTREVUE AVEC MONSIEUR FRANCOIS REY (cf. p.10)	p.35
4) EXEMPLES ET EXPLICATIONS DU CALCUL DE LA DISTANCE INTERTEXTUELLE (cf. p.13)	p.36
5) RESULTATS DE LABBE SUR LES TEXTES GARY-AJAR COMPLET (cf. p.28)	p.39
<u>BIBLIOGRAPHIE</u>	p.41
<u>REMERCIEMENTS</u>	p.42

INTRODUCTION

Le langage et l'écriture sont des facultés universelles, il n'empêche que chaque personne a sa propre manière de s'exprimer, à force de mots, d'expressions, de ponctuations... Il en est de même pour les écrivains de quelque époque que ce soit, et c'est cette façon personnelle de rédiger qui est, en quelque sorte, leur marque de fabrique. Ce TPE tentera de répondre à la question : une étude statistique peut-elle permettre d'attribuer ou non une œuvre à un auteur ?

En effet, il y a plus de cent ans a été inventée une technique méconnue, mais révolutionnaire : la stylométrie ou l'étude statistique des textes. La stylométrie désigne un ensemble de méthodes statistiques qui ont pour but d'apporter des éclairages à la stylistique. Le présupposé de cette discipline est qu'il y aurait des traits quantifiables qui signeraient la marque d'un auteur, d'un texte, d'un genre. Cet examen est généralement basé sur une étude du lexique.

T. Mendenhall fut le pionnier du genre : il appliqua cette technique, d'une façon néanmoins très rudimentaire, en 1887, sur un texte attribué à Shakespeare, par certains, et à Francis Bacon, par d'autres. Sa méthode consistait à étudier la longueur des mots. La conclusion fut que l'auteur de ce récit était bel et bien Shakespeare (il est à noter que Shakespeare est le centre de différents questionnements, dont une incertitude sur son existence même).

Mais depuis, la stylométrie s'est nettement améliorée, notamment depuis les travaux de Yule et Zipf en 1944 sur la modélisation de la distribution du lexique dans les textes ; il faut également citer l'imposant travail de Mosteller et Wallace en 1964 sur l'attribution des "Federalist papers"* qui constitue la première véritable grande étude statistique de textes. Grâce à ces deux faits, la stylométrie est devenue aux yeux de nombreux mathématiciens une preuve pour l'attribution certaine d'une œuvre à un auteur. Elle a permis ainsi de soulever des débats autrement importants et passionnants. Celui que nous exposerons dans la majeure partie de notre TPE concerne deux des écrivains les plus célèbres de la littérature française, qui ont marqué le dix-septième siècle : et si Corneille avait écrit la plupart des œuvres de Molière ?

Cette question n'est cependant pas née du hasard, elle a été posée par plusieurs personnes ayant étudié les œuvres, devenues classiques, de ces deux auteurs ; et c'est pourquoi, avant de vous proposer l'analyse statistique effectuée sur leurs œuvres et ce qui l'entoure, notre première partie portera sur ce que l'on appelle « L'affaire Corneille-Molière », d'un point de vue purement biographique et présentant principalement les arguments de ceux qui défendent cette thèse. Enfin, nous terminerons par une étude faite par nos soins, au sujet de "deux" autres auteurs dont la conclusion est connue et sûre (Gary et Ajar).



CORNEILLE ET MOLIÈRE

1) Pierre Corneille et Jean-Baptiste Poquelin

Tout d'abord, nous allons vous présenter les deux protagonistes de l'affaire en proposant pour chacun leur biographie officielle.



Pierre Corneille est né le 6 juin 1606 à Rouen. Issu d'une famille de magistrats, il accomplit des études secondaires au Collège des Jésuites de Rouen avant de s'orienter vers une carrière d'avocat. En 1629, un chagrin amoureux le conduit à écrire sa première comédie, *Mélite ou les fausses lettres*.

Face au succès fulgurant de cette pièce, Corneille décide d'entreprendre une carrière d'auteur de théâtre. De 1630 à 1636, il écrit cinq nouvelles comédies dont *L'Illusion comique* : ces dernières contribuent à redonner ses lettres de noblesses à un genre jugé secondaire auquel on reproche sa vulgarité et ses outrances ; en effet, Corneille a su utiliser une voie qui refuse le grotesque pour privilégier la peinture des caractères et des mœurs.

En 1635, il écrit *Médée*, sa première tragédie. Suite au succès de ses différentes pièces, Corneille est engagé la même année par Richelieu, avec quatre autres auteurs, pour mettre en vers les intrigues théâtrales que le cardinal se plaît à imaginer. En janvier 1637, Corneille fait jouer *Le Cid*, tragi-comédie inspirée d'une épopée espagnole, au Théâtre du Marais : la pièce connaît un immense triomphe. Au printemps de l'année 1641, il épouse Marie de Lampérière avec laquelle il aura sept enfants. En 1642, la mise en scène de *Cinna ou la Clémence d'Auguste* lui apporte la consécration : il apparaît alors comme le plus grand poète dramatique de son époque et on le surnomme le "Sophocle français". Célébré par le public, reconnu par ses pairs, financé par le pouvoir, il connaît une décennie éclatante. En 1647, il entre à l'Académie française.

Suite à l'échec de *Pertharite* en 1651, Corneille décide de renoncer au théâtre, se considérant trop "vieux" pour écrire. Ce n'est qu'en 1659 que Corneille décide de renouer avec la tragédie en écrivant *Œdipe* : cette pièce connaît un succès colossal. Suite au succès triomphal de Racine avec *Andromaque*, en 1667, Corneille voit apparaître en ce dernier un rival à sa hauteur : en 1670, les deux auteurs se trouvent en concurrence frontale lorsqu'ils créent simultanément une pièce sur le même thème ; Racine triomphe avec sa *Bérénice*, tandis que le *Tite et Bérénice* de Corneille ne rencontre qu'un succès mitigé.

En 1672, Corneille fait jouer *Pulchérie*, une comédie héroïque : c'est un échec suivi deux ans plus tard de celui de *Suréna*, ce qui pousse Corneille à cesser définitivement son activité de dramaturge en 1674. Dix ans plus tard, le 1er octobre 1684, il décède à Paris laissant derrière lui un immense héritage théâtral.

Jean-Baptiste Poquelin, né à Paris en 1622 dans la famille des tapissiers du roi, étudie au collège jésuite de Clermont (lycée Louis le Grand) où il est en classe avec le prince de Conti, qui deviendra l'un de ses protecteurs, et où il apprend le latin, le grec, la rhétorique et les dogmes de l'Église. Il débute des études de droit en 1640 et sera reçu en 41.

En 1643, contre l'avis familial, il décide de devenir comédien, créant une troupe, « L'Illustre Théâtre », avec sa maîtresse de l'époque, Madeleine Béjart, la fille de celle-ci et quelques autres comédiens. Il prend le nom de Molière en 1644. Les raisons qui l'ont incité à choisir ce pseudonyme n'ont jamais été élucidées. Après des débuts difficiles et une tournée de treize ans à travers la France, la troupe, dirigée désormais par Molière, s'installe à Paris en 1658, suite à un séjour à Rouen.

Elle se produit alors devant le roi et reçoit un soutien important, celui de Monsieur, son frère. Grâce à ce dernier, la compagnie s'installe, et ce définitivement, au théâtre du Palais-Royal ; Molière, lui, épouse en 1662 Armande Béjart, fille ou sœur de Madeleine, et se lance dans la rédaction de pièces dont *Sganarelle ou Le Cocu imaginaire* (1660), *Les Fâcheux* (1661) ou *L'école des femmes* (1662) ; toutes remportent un vif succès.

En 1664, sa pièce *Le Tartuffe* est interdite sous la pression du clergé de même que *Dom Juan* (1665), mais la troupe devient « Troupe du Roi » la même année. Après des années calmes sur le plan général, mais riches au niveau de la création théâtrale (*Le Misanthrope*, 1666), *Amphitryon* (1668), *Le Bourgeois Gentilhomme* (1670), *Les Fourberies de Scapin* (1671), *Les Femmes savantes* (1672), etc.), Jean-Baptiste Poquelin, dit Molière, s'éteint le 17 février de l'an 1673, au sortir d'une représentation de la pièce du *Malade Imaginaire*.

Ces deux auteurs ont marqué le 17^{ème} siècle, l'un par ses comédies l'autre par ses tragédies, mais, au cours des siècles, une polémique s'est ouverte : et si toutes ces célèbres pièces n'étaient le fruit du génie que d'un seul auteur : Corneille ?



2) Les débuts de l'Affaire et son évolution

L'affaire commença réellement à la fin de l'année 1919, lorsque le poète et écrivain Pierre Louÿs, après plusieurs années d'études stylistiques, tenta de publier la preuve que Corneille était l'auteur des plus grandes pièces de Molière, telles *Les Femmes savantes*, *Dom Juan* ou *Amphitryon*... Il ne put jamais y parvenir, face à la grogne et aux menaces de ceux qui se pensaient experts, moliéristes* ou dix-septiémistes*, laissant dans l'oubli les centaines de pages qu'il avait rédigées.

Pierre Louÿs sera certes méprisé, insulté, mais sa thèse n'a pas laissé indifférent tout le monde : quelques années plus tard elle est reprise par une universitaire anglaise dans plusieurs articles et un ouvrage, sans réel impact.

En 1951, Henry Poulaille affirme que *Le Tartuffe* a été écrit par Corneille ; du même auteur est publié en 1957 *Corneille sous le masque de Molière*, et sa thèse est soutenue également par l'écrivain René-Louis Doyon.

Mais après 1958 l'affaire tombe dans une phase d'oubli. Il faudra attendre 1990 pour qu'elle réapparaisse avec le livre du Belge Hippolyte Wouters, *Molière ou l'auteur imaginaire ?*, qui restera néanmoins peu convaincant.

2001 constitue une année clef pour cette affaire avec la première publication de l'étude statistique réalisée par Dominique Labbé et exposée dans son livre *Corneille dans l'ombre de Molière*, en 2003. Depuis, le chef de file des cornéliens* est Denis Boissier qui a publié en 2004 *L'affaire Molière*, puis en 2007 une thèse de 1000 pages *Molière Bouffon du Roi et prête-nom de Corneille*.

Mais ce qui a véritablement changé depuis l'époque de Pierre Louÿs, outre de nouveaux faits historiques et scientifiques apportés, est le récent engouement de la presse pour cette affaire : en effet,

les journaux s'y sont intéressés comme le montrent les deux récentes pages parues dans *l'Express*, mais aussi une émission sur France 2 qui a présenté l'affaire dans la série de Stéphane Bern *Secrets d'histoire : Molière a-t-il écrit ses pièces ?*

3) Arguments biographiques et stylistiques

Ces différentes personnes ont recensé dans les biographies admises comme officielles de nombreuses incertitudes quant aux faits présentés, des similitudes étranges entre la vie des deux personnages ainsi que des problèmes non élucidés. C'est pourquoi cette partie, en reprenant les arguments des cornéliens, va répondre à trois questions : Pourquoi Molière n'aurait-il pas écrit ses pièces ? Pourquoi aurait-ce été Corneille ? Quels avantages auraient-ils pu tirer de cette situation ?

Molière est issu de la famille des tapissiers du roi : son père, qui a acheté cette charge, était tapissier, son grand père l'était, ainsi que son arrière grand-père... Or à cette époque, la majeure partie de la population était illettrée et n'allait donc pas à l'école, de ce fait pourquoi un fils de tapissier y serait-il allé ? Voltaire affirme en 1739 dans sa *Vie de Molière*, dès la première page : « le père marchand fripier et la mère lui donnèrent une éducation trop conforme à leur état, auquel ils le destinaient: il resta jusqu'à quatorze ans dans leur boutique, n'ayant rien appris, outre son métier, qu'un peu lire et écrire ». Mais il est bien sûr impensable que l'un des plus grands auteurs français n'ait pas fait de solides études. Les moliéristes ont donc affirmé que Molière a étudié dans le collège des Jésuites de Clermont, alors qu'aucun papier n'atteste son inscription. Les moliéristes soutiennent également que Molière a fait de brillantes études en compagnie du prince de Conti qui, après vérification, est de sept ans son cadet !

De même, les études de droit de Molière ne sont absolument pas prouvées. Dans ses pièces on en retrouve beaucoup de notions, on a donc déduit que Molière avait fait de telles études, mais, encore une fois, aucune trace de son inscription. A contrario, Corneille fut licencié à l'âge de dix-huit ans...

Éperdu d'amour pour la belle Madeleine Béjart, Molière s'engage dans la troupe de celle-ci, qui l'emmène jusqu'à Rouen, où elle séjourne en 1643 ; à cette occasion, Jean-Baptiste rencontre pour la première fois l'homme qui marquera sa vie : le dramaturge Corneille ; c'est alors qu'il repart avec son célèbre pseudonyme, Molière. S'en suit son tour de France d'une douzaine d'années, durant lesquelles il n'écrira aucune grande pièce, faute d'imagination ?

Cependant, Poquelin commence à publier des œuvres à partir de 1658. Il avait, certes, écrit deux pièces, *L'Etourdi* et *Le Dépit amoureux*, qu'il avait présentées, à Lyon en 1655. Il semble cependant qu'elles aient été réécrites pour créer une version dite « parisienne » (cf. annexe 1, page 34). En 1658 également, il fait avec sa troupe une seconde étape dans la ville de Rouen. Il se rapproche ainsi à nouveau de Corneille, avec lequel on soupçonne qu'il ait passé un accord... Les deux personnages partent alors pour Paris, à quelque temps d'intervalle, où la troupe du plus jeune, alors inconnue, est présentée directement au frère du roi, devant lequel elle joue la tragédie *Nicomède* de Corneille, suivie du *Médecin Volant* de Molière inspirée très fortement d'une pièce italienne de l'époque (*Il medico volante*).

De plus, au cours de sa nouvelle vie, Molière acquiert toutes les caractéristiques d'un bouffon du roi et de nombreuses charges. Il est en effet le directeur du plus grand théâtre de Paris et s'attribue les rôles les plus importants !

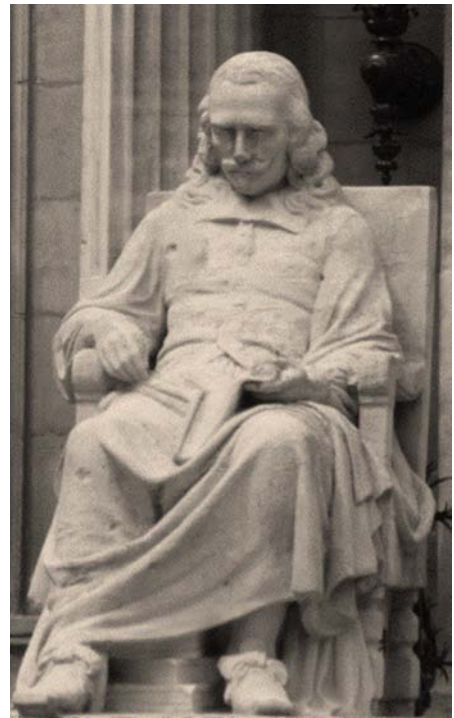
Mais, mis à part son manque de temps, ses études douteuses, un autre fait étrange est, outre qu'il n'ait jamais signé ses pièces, que l'on n'a retrouvé aucun écrit, aucune note, aucun document lui appartenant.

Enfin, dernier fait troublant, le style de Molière : ce dernier est le seul à arriver à s'approcher de celui de Corneille ; même de grands auteurs, dont le propre frère de Pierre, Thomas, n'ont pas réus-

si. Racine affirme que le style de Corneille est « inimitable » comme le montre son hommage adressé à celui-ci, présenté dans l'annexe 2, page 36.



Statues de Molière et Corneille,
devant l'opéra-théâtre d'Avignon



Voici, à présent, une comparaison réalisée par nos soins de deux extraits de textes dont le style paraît semblable :

Texte I

« Quoi ! Tu ne **me** dis mot ! Crois-tu que ton **silence**
Puisse de tes **discours réparer l'insolence** ?
Des **pleurs** effacent-ils un **mépris si cuisant**,
Et ne t'en dédis-tu, **traître**, qu'en te **laissant** ?
Pour **triompher de moi**, veux-tu, pour toutes **armes**,
Employer des **soupirs** et de **muettes larmes** ?
Sur **notre amour passé** c'est trop te **confier**,
Dis du moins quelque chose à te **justifier**,
Demande le **pardon** que tes **regards m'arrachent**.
Explique leurs **discours**, dis-**moi** ce qu'ils **me cachent**. »

Texte II

« Ah ! Que vous savez bien ici, contre **moi-même**,
Perfide, vous **servir de ma faiblesse extrême**,
Et ménager pour vous l'**excès prodigieux**
De ce **fatal amour** né de vos **traîtres yeux** !
Défendez-vous au moins d'un **crime qui m'accable**,
Et cessez d'**affecter** d'être avec **moi coupable**.
Rendez-moi, s'il se peut, ce **billet innocent** ;
A vous **prêter les mains ma tendresse consent** ;
Efforcez-vous ici de paraître **fidèle**,
Et je **m'efforcerai, moi**, de vous **croire** telle. »

Après une simple lecture de ces deux textes, nous ne pouvons déterminer avec certitude l'auteur de chacun de ces deux extraits. S'agit-il de deux textes de Molière ? De deux textes de Corneille ? Ou encore de textes de Corneille et de Molière ?

En effet, ces deux textes semblent très proches au niveau du style et de la tonalité : les interjections « Quoi ! » et « Ah ! » annoncent dès le début un ton dramatique, voire grave. Nous pouvons constater une omniprésence du champ lexical de la passion et du drame (voir mots surlignés en rouges), termes récurrents dans les tragédies. Ces mots sont placés dans les deux textes dans un ordre logique aggravant ainsi le côté tragique du discours, ils sont aussi accentués par des adjectifs et des prépositions (voir mots surlignés en bleu).

Outre une omniprésence des champs lexicaux vue précédemment, nous remarquons une surabondance d'adjectifs possessifs et de pronoms compléments à la première personne du singulier et du pluriel (voir mots surlignés en jaune) utilisé par le locuteur afin d'exprimer ses sentiments personnels : il s'agit de lyrisme.

Nous pouvons aussi constater que ces deux extraits ont une structure commune : ils sont en vers, chaque vers étant lui-même un alexandrin. La répétition des mots « amour » et « traître » annonce aussi la thématique du discours.

Alors, à votre avis, s'agit-il de deux textes de Corneille ? Non, un texte de Molière et un autre de Corneille ? Mais lequel est de Molière ?

Il s'agit, en fait, pour le texte 1 d'un extrait de *La Place royale* (acte III, scène 6) de Corneille tandis que le texte 2 est extrait du *Misanthrope* (acte IV, scène 3) de Molière.

En plus de nombreux traits techniques similaires, le contenu est lui aussi très proche : ce sont deux personnes dépitées qui parlent, blessées par la tromperie de leur amant. Angélique (texte 1) et Alceste (texte 2) se croient trompés et souhaitent ardemment qu'Alidor et Célimène se justifient de leurs méfaits pour mieux les pardonner :

-Dans *La Place Royale* de Corneille :

«Dis du moins quelque chose à te justifier,
Demande le pardon que tes regards m'arrachent.
Explique leurs discours, dis-moi ce qu'ils me cachent »

-Dans *le Misanthrope* de Molière :

«Défendez-vous au moins d'un crime qui m'accable, »
« Efforcez-vous ici de paraître fidèle,
Et je m'efforcerai, moi, de vous croire telle. »

Avec l'utilisation de l'impératif, leur souhait se transforme presque en ordre. Mais les deux accusés ont la même réaction : ils ne se défendent pas, ne parlent pas :

« Et ne t'en dédis-tu, traître, qu'en te taisant ? »

« Défendez-vous au moins d'un crime qui m'accable, »

Même si le texte 1 présente un plus grand champ lexical de la parole que le second, on devine aisément la réaction de Célimène.

Après avoir vu ces similitudes étranges du point de vue biographique et stylistique, les moliéristes se sont demandés quels avantages aurait pu tirer Corneille d'une telle collaboration. Cinq avantages majeurs se dégagent selon Denis Boissier :

- tout d'abord, grâce à ce prête-nom, Corneille peut renouer avec la comédie de mœurs qu'il avait dû abandonner après le succès de ses tragédies notamment *Le Cid*, joué pour la première fois en 1636 ; il était en effet, à cette époque, très mal vu d'écrire des farces ;

- grâce à ce prête-nom, Corneille n'aurait plus eu à redouter les foudres de l'Église et celles de la Sorbonne, c'est à dire de l'élite, qui n'avait cessé de critiquer ses tragédies depuis *Polyeucte* (1642) ;

- il ne faut pas oublier que Molière était vraisemblablement bouffon du roi ; écrire pour Molière rapprocherait donc Corneille de sa Majesté Louis XIV, et lui permettrait de travailler officieusement pour lui et ainsi passer outre les critiques ;

- les relations qu'il avait avec Molière le rapprocheraient de la troupe et donc lui permettraient de rester proche de certaines femmes qu'il chérira longtemps comme Marquise du Parc, ou encore Armande Béjart épouse de Molière ;

- reste l'argent. Effectivement, Corneille aura, durant toute sa vie, de nombreux problèmes d'ordre pécuniaire, dus principalement à ses devoirs de père - à savoir pourvoir à l'éducation, puis aux études de ses six enfants : ses deux fils aînés embrassant une carrière militaire onéreuse – or, de l'argent, cette association lui en aurait procuré abondamment. Molière dirige le Palais-Royal dès 1661, et l'auteur qu'il interprète et qu'il interprétera le plus (mis à part lui-même) n'est autre que Corneille ; le comédien jouera au moins neuf pièces du dramaturge, ce dernier en écrira officiellement trois pour la troupe (*Tite et Bérénice*, *Psyché*, *Pulchérie*). Pour chacune, succès ou non, Corneille touchera 2 000 livres, ce qui reste la somme la plus importante qu'ait jamais perçue un auteur de l'époque.

Afin de ne pas alourdir notre exposé, nous avons choisi de ne pas nous étendre sur les arguments des adversaires de cette thèse : ils sont en effet, pour la plupart, simplistes.

Nous avons cependant rencontré Monsieur François Rey, grand connaisseur et défenseur de Molière, auteur de *Molière et le roi*, maître de conférences, qui a eu l'amabilité de nous recevoir et de répondre à nos questions. Nous l'avons interrogé sur l'éducation de Molière, sa rencontre avec Corneille, son manque de temps pour écrire, son surnom et sur l'étude statistique de M. Labbé. Nous vous proposons donc de vous rendre à l'annexe 3 (page 35) où vous trouverez un résumé de cette entrevue.

4) Conclusions

Ces différents faits et arguments montrent, malgré tout ce que l'on peut penser de l'affaire, de véritables incertitudes et incompréhensions au sujet de Molière, de sa vie et de son œuvre : la thèse des cornéliens paraît réellement fondée et même viable.

Cependant, en l'absence d'écrit de Molière ou de toute note pouvant affirmer que Corneille a bien aidé son ami en lui écrivant des pièces et en se servant de lui comme d'un prête-nom, l'affaire ne peut que rester au *statu quo*.

Toutefois, les mathématiques se sont mêlées à cette histoire, apportant de nouveaux arguments par l'étude statistique des textes des deux hommes et soutenant ainsi la position cornélienne. C'est ce que nous allons maintenant voir dans la partie suivante qui présentera, entre autres, les résultats de l'analyse de celui qui a mené cette expérience hors du commun.



Le poète Pierre Louÿs,
premier cornélien

L'AFFAIRE ET LES MATHÉMATIQUES

1) La stylométrie

Comme nous l'avons présenté en introduction de notre dossier, la stylométrie fut inventée voilà maintenant plus d'un siècle. Elle n'a, dès lors, cessé de s'améliorer, principalement depuis l'avènement de l'informatique, du scanner et des logiciels créés à cet effet. Elle est, aujourd'hui, une branche à part entière de la statistique, malgré la réticence de nombreux littéraires qui ne veulent pas voir les mathématiques s'immiscer dans leur domaine, phénomène que l'on retrouve surtout en France... Cependant, peut-être le résultat de telles expériences est-il plus concluant qu'ils ne veulent bien le croire, l'ordinateur n'exécutant que des calculs absents de tout préjugé.

Utilisée principalement sur les discours politiques et les textes historiques ou littéraires, la stylométrie est employée à différentes fins. Par exemple, dater l'un d'eux, définir le style d'un auteur auquel l'on pourra rattacher des œuvres, comme dans cette affaire. Mais elle peut également mettre en évidence des thèmes volontairement camouflés. Elle peut être réalisée de différentes façons, souvent combinées pour confirmer le résultat trouvé et éviter les critiques de ceux qui douteraient encore plus, si l'on en avait utilisé qu'une seule.

L'on trouve, parmi elles, différents types de calculs, basés entre autres sur la longueur des mots ou des phrases, le nombre de syllabes par vocable*, la distance intertextuelle, la fréquence des combinaisons de mots ou leur champ sémantique... Ces trois dernières techniques sont celles auxquelles nous allons nous intéresser, car utilisées dans l'étude comparative des œuvres de Corneille et de Molière réalisée par Monsieur Dominique Labbé.



2) L'étude de Dominique Labbé

Dominique Labbé, docteur en sciences politiques, maître de conférences à l'Institut d'études politiques de Grenoble et spécialiste de l'analyse du discours, développe depuis 1969 des outils statistiques et informatiques pour l'étude du discours. Appliqués d'abord sur les discours politiques de l'époque (de Gaulle, Mitterrand...), ils se sont révélés adaptés à tout texte en français et donc à l'ensemble de la littérature de ces quatre derniers siècles, chose rendue possible, à grande échelle, par la numérisation.

C'est ainsi que D. Labbé s'est rendu compte du rapport évident entre les textes de Molière et de Corneille - la similitude qu'il a remarquée ne se retrouve qu'entre les textes de Romain Gary et d'Emile Ajar -, entraînant en 2003 la publication de *Corneille dans l'ombre de Molière*, où il expose sa technique et ses résultats, et *Si Deux et deux sont quatre...* en 2009, ouvrage plus historique et biographique.

a) Le calcul de la distance intertextuelle

Cette technique, la première utilisée par Dominique Labbé, est sans doute la plus importante, la plus fastidieuse et celle faisant le plus appel aux mathématiques.

Voici donc, d'abord, une explication sommaire et générale, accompagnée d'une démonstration mathématique et suivie des résultats et conclusions de l'étude appliquée aux textes de Molière et de Corneille.

- La distance entre deux textes se mesure comme la distance séparant deux objets dans l'espace. L'unité de mesure est le "mot". Soit deux textes A et B :



La distance intertextuelle s'obtient en superposant ces deux textes et en comptant le nombre de mots différents. Elle est nulle si les deux textes se recouvrent totalement et est égale à la taille des deux textes s'ils n'ont aucun mot en commun.

A et B confondus :
la distance est nulle



A et B disjoints : la
distance est égale au
total des mots utilisés
dans les deux textes



Pour tous les vocables appartenant à A et/ou B, on calcule la différence absolue de leurs effectifs dans A et B et on fait la somme de ces différences.

Afin de rendre comparable ces distances quand elles sont calculées sur des textes de longueurs différentes, on divise cette somme par le nombre total de mots compris dans A et B. On obtient un indice qui varie uniformément entre 0 (tous les vocables ont les mêmes effectifs dans A et B) et 1 (aucun vocable commun entre A et B).

On détermine ainsi différentes échelles : en dessous de 0.20, on affirme que le texte est bien du même auteur ; entre 0.20 et 0.25, l'auteur est très probablement le même et si il est prouvé qu'il s'agit de deux différents, il s'agit soit d'une collaboration soit du plagiat de l'un sur l'autre ; pour une distance comprise entre 0.25 et 0.35, il s'agit principalement de textes d'un même auteur traitant de sujets différents, ou de deux traitant du même ; au-delà, il est à peu près certain que les auteurs sont différents.

Pour les expériences d'attribution d'auteur, il faut bien sûr confronter des textes contemporains écrits dans un même genre littéraire. Ceux qui sont séparés par les distances les plus faibles sont d'un même auteur.

Calcul :

Soit V un vocable*, et A et B deux textes.

V_a = nombre de V dans A

M_a = nombre de mots dans A

V_b = nombre de V dans B

M_b = nombre de mots dans B

Pour tous les vocables appartenant à A, B on calcule :

Soit V un vocable*, et A et B deux textes.

V_a = nombre de V dans A

M_a = nombre de mots dans A

V_b = nombre de V dans B

M_b = nombre de mots dans B

Pour tous les vocables appartenant à A, B on calcule :

$$\text{On calcule : } \frac{\sum_v |V_a - V_b|}{M_a + M_b} = \text{indice} = \text{distance intertextuelle}$$

Si les textes sont identiques, $V_a - V_b = 0$ pour tout V , donc l'indice est de 0.

Si les textes sont disjoints (aucun vocable en commun),

$$|V_a - V_b| = V_a \text{ ou } V_b$$

$$\sum_v |V_a - V_b| = \sum_v V_a + \sum_v V_b = M_a + M_b$$

$$\frac{M_a + M_b}{M_a + M_b} = 1$$

Exemple

:

Voici un court extrait de 20 mots choisis au hasard dans le *Notre Dame de Paris* de Hugo : « La chèvre le regarda d'un œil intelligent, leva son pied doré et frappa sept coups. Il était sept heures. ». Nous l'avons comparé avec une réécriture dans laquelle seuls quelques termes ont été changés et avec un texte de la même longueur qui n'a que peu de termes en commun.

Nous vous proposons de vous rendre à l'annexe 4, page 38, où vous trouverez explications et résultats trouvés.

Voici le tableau récapitulatif des résultats lors des comparaisons des pièces de Molière avec les deux *Menteurs* de Corneille et *Les Plaideurs* de Racine, cette dernière est présente pour appuyer les résultats et montrer qu'ils ne sont pas seulement proches à cause du style ou de l'époque.


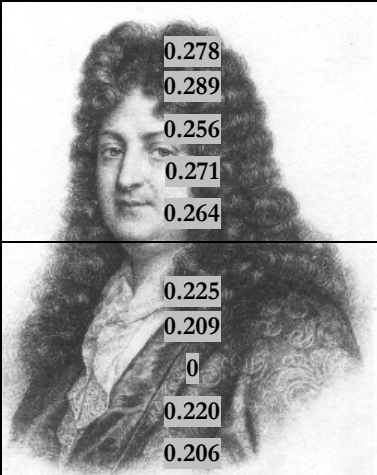
N°	Pièces	Genre	Le menteur (Corneille 1642)	Suite du menteur (Corneille 1643)	Les plaideurs (Racine : 1668)
15	Le menteur (1642)	Vers	0,000	0,180	0,296
16	La suite du menteur (1643)	Vers	0,180	0,000	0,293
34	Psyché Corneille (1671)	Vers	0,288	0,273	0,348
36	Psyché Molière (1671)	Vers	0,329	0,325	0,354
37	La Jalousie du barbouillé (avant 1660)	Prose	0,341	0,331	0,327
38	Médecin volant (avant 1660)	Prose	0,310	0,293	0,302
39	L'Etourdi (1658)	Vers	0,205	0,206	0,269
40	Dépit amoureux (1658)	Vers	0,215	0,212	0,270
41	Précieuses ridicules (1660)	Prose	0,315	0,314	0,314
42	Sganarelle ou le cocu imaginaire (1660)	Vers	0,259	0,253	0,293
43	Dom Garcie de Navarre (1661)	Vers	0,280	0,273	0,359
44	L'École des maris (1661)	Vers	0,223	0,217	0,279
45	Les Fâcheux (1661)	Vers	0,248	0,248	0,306
46	L'École des femmes (1662)	Vers	0,226	0,217	0,261
47	Critique de l'école des femmes (1663)	Prose	0,323	0,319	0,340
48	L'Impromptu de Versailles (1663)	Prose	0,321	0,316	0,323
49	Mariage forcé (1664)	Prose	0,322	0,302	0,320
50	Princesse d'Elide (1664)	Vers Prose	0,251	0,243	0,314
51	Le Tartuffe (1664)	Vers	0,242	0,228	0,275
52	Dom Juan (1665)	Prose	0,259	0,248	0,281
53	L'Amour médecin (1665)	Prose	0,292	0,289	0,287
54	Le Misanthrope (1666)	Vers	0,252	0,234	0,283
55	Médecin malgré lui (1666)	Prose	0,298	0,289	0,296
56	Mélicerte (1666)	Vers	0,257	0,250	0,322
57	Le Sicilien ou l'amour peintre (1667)	Prose	0,277	0,260	0,301
58	Amphytrion (1668)	Vers libres	0,253	0,256	0,297
59	Georges Dandin (1668)	Prose	0,292	0,279	0,292
60	L'Avare (1668)	Prose	0,256	0,244	0,270
61	M. de Pourceaugnac (1669)	Prose	0,292	0,283	0,285
62	Amants magnifiques (1670)	Prose	0,282	0,279	0,329
63	Bourgeois gentilhomme (1670)	Prose	0,294	0,280	0,286
64	Fourberies de Scapin (1671)	Prose	0,269	0,263	0,281
65	Comtesse d'Escarbagnas (1671)	Prose	0,311	0,300	0,305
66	Femmes savantes (1672)	Vers	0,260	0,248	0,283
67	Malade imaginaire (1672)	Prose	0,282	0,270	0,278
<i>Moyenne oeuvre de Molière</i>			0,275	0,266	0,299
<i>Moyenne pièces en vers de Molière</i>			0,241	0,234	0,290
<i>Moyenne oeuvre de Corneille</i>			0,252	0,249	0,347
<i>Moyenne oeuvre de Racine</i>			0,314	0,311	0,376



Distances séparant *Dom Garcie* (Molière) et *Psyché* (Corneille et Molière) des dernières pièces de Corneille, montrant la promiscuité du style de Molière de celui de Corneille...

Ultimes pièces de Corneille	Dom Garcie (Molière,1661)	Psyché (Corneille 1671)
Rodogune (1644)	0,245	0,231
Theodore (1645)	0,234	0,245
Heraclius (1647)	0,248	0,273
Andromède (1650)	0,241	0,218
Don Sanche (1650)	0,224	0,251
Nicomède (1651)	0,244	0,264
Pertharite (1651)	0,235	0,263
Edipe (1659)	0,223	0,226
Toison d'or (1661)	0,221	0,220
Sertorius (1662)	0,230	0,238
Sophonisbe (1663)	0,228	0,236
Othon (1664)	0,235	0,240
Agesilas (1666)	0,234	0,233
Attila (1667)	0,235	0,227
Tite et Bérénice (1670)	0,227	0,235
Psyché (1671)	0,230	—
Pulcherie (1672)	0,230	0,226
Surena (1674)	0,216	0,224
<i>Moyenne Corneille</i>	<i>0,243</i>	<i>0,244</i>
<i>Moyenne Molière</i>	<i>0,286</i>	<i>0,297</i>

... et distances entre les dernières pièces de Corneille et celles contemporaines de Racine.

	Tite et Bérénice (Corneille, 1670)	Bérénice (Racine, 1670)
CORNEILLE :		
Agésilas (1666)	0.159	0.278
Attila (1667)	0.180	0.289
Tite et Bérénice (1670)	0	0.256
Pulchérie (1672)	0.155	0.271
Suréna (1672)	0.156	0.264
RACINE :		
Andromaque (1667)	0.259	0.225
Britannicus (1669)	0.251	0.209
Bérénice (1670)	0.256	0
Bazajet (1672)	0.262	0.220
Mithridate (1673)	0.249	0.206

Dans chacun des tableaux, les résultats les plus troublants sont écrits en gras. Les œuvres dont le titre est écrit de la même façon sont celles pour lesquelles Dominique Labbé n'a aucun doute lorsqu'il affirme que Corneille en est l'auteur.

b) Les combinaisons de mots

Cette seconde phase fait appel à des moyens de classification, et plus particulièrement à la lemmatisation, dont voici une explication :

- Ce terme désigne l'analyse lexicale du contenu d'un texte qui regroupe des mots d'une même famille. Chacun des mots se trouve ainsi réduit en une entité, appelée lemme ou forme canonique : la lemmatisation regroupe les différentes formes que peut revêtir un mot (nom singulier, nom pluriel, verbe à l'infinitif ou conjugué...).

La forme canonique est : - pour un verbe, sa forme infinitive.

- pour les autres, le mot masculin singulier (sauf exception).

Exemples : - l'adjectif gentil existe sous quatre formes : gentil, gentille, gentils et gentilles : la forme canonique de tous ces mots est « gentil ». - le verbe « avoir » possède, lui, des multitudes de formes : ai, as, a, avaient, aurez, ayant, eussions eu... La forme canonique est « avoir » pour chacun.

Après cette étape, que Dominique Labbé a appliquée aux combinaisons « pseudo-auxiliaire + infinitif », on peut retrouver les plus fréquemment employées dans les œuvres de Corneille, Molière et Racine.

En comparant les résultats, cela fait apparaître que Racine ne partage que trois combinaisons principales avec ses compères et que celles-ci sont dans un ordre et dans des proportions très éloignées. (F représente la fréquence pour 100 000 mots)

P. Corneille		Molière		Racine	
Syntagmes	F	Syntagmes	F	Syntagmes	F
<i>faire voir</i>	33,8	<i>faire voir</i>	31,5	aller voir	12,0
<u><i>pouvoir être</i></u>	18,8	<u><i>pouvoir être</i></u>	25,5	<u>pouvoir voir</u>	9,6
<u><i>pouvoir faire</i></u>	18,4	<u><i>pouvoir faire</i></u>	25,5	faire entendre	9,0
faire naître	13,9	vouloir dire	24,9	<u>pouvoir faire</u>	8,4
<u><i>pouvoir voir</i></u>	13,4	<i>vouloir faire</i>	19,5	aller chercher	7,8
devoir être	12,7	pouvoir dire	14,5	faire parler	7,8
pouvoir souffrir	10,8	pouvoir avoir	13,7	<u>pouvoir être</u>	7,8
<i>vouloir faire</i>	9,9	aller faire	13,2	venir chercher	7,2
faire connaître	9,6	avoir faire	13,2	faire éclater	6,6
devoir faire	8,7	<u><i>pouvoir voir</i></u>	12,3	falloir partir	6,6

En revanche, Corneille et Molière en partagent cinq, dont trois majoritaires qui arrivent dans le même ordre et avec des densités comparables : pour Dominique Labbé, ce fait est encore une fois totalement

unique dans l'histoire littéraire française, mis à part l'affaire Gary-Ajar (les termes soulignés sont ceux que l'on retrouve chez les trois auteurs).

c) *Le sens des mots*

Ce dernier modèle entrant dans le travail de lexicométrie de Dominique Labbé peut, de nom, sembler étrange du fait de sa réalisation informatique.

Il est lui aussi effectué avec l'aide de la lemmatisation et de procédés de classification, qui cherchent, dans les environs d'un lemme, les mots qui lui sont les plus souvent associés.

Pour le chercheur, le terme le plus évocateur est le mot « amour » : le logiciel dresse la liste de toutes les formes canoniques et compte celles qui sont le plus souvent rattachées au mot type, ainsi on trouve :

- que l'amour est très souvent accompagné des verbes « céder », « étreindre », « opposer », « allumer », « naître », etc.

- que la « haine » est le terme le plus couramment employé avec l'amour.

Remarques : - le professeur pose tout de même une limite à cette méthode appliquée aux textes versifiés tels qu'ici : toujours dans cet exemple de l'amour, arrivent à la suite de la haine des mots tels que « tour » (masculin) ou « jour » qui sont présents pour la rime.

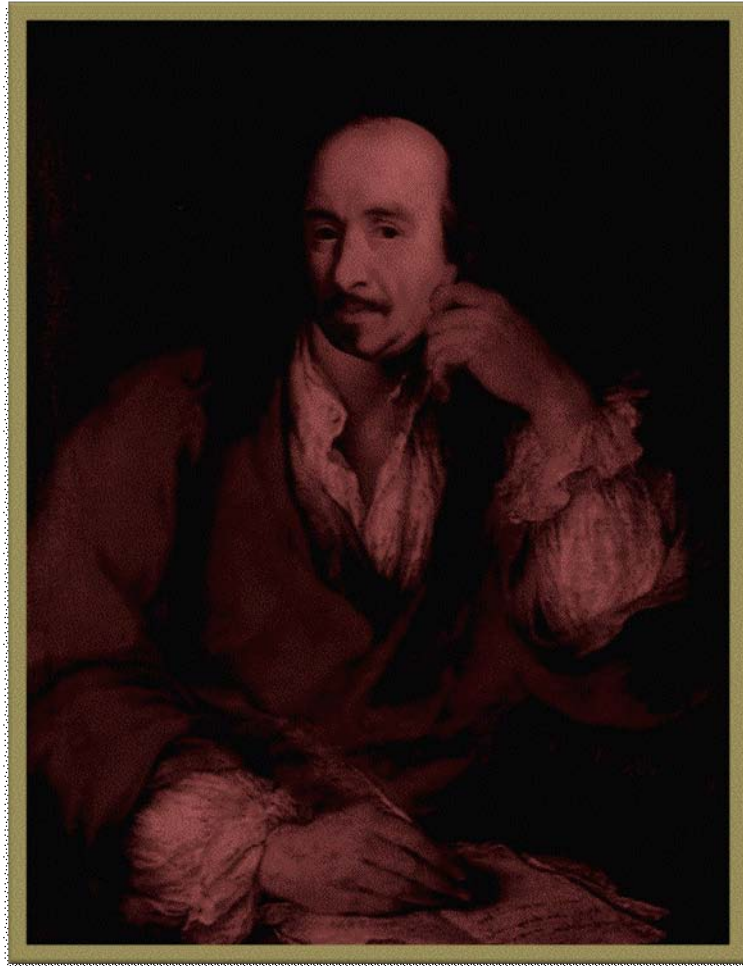
- cette méthode est aussi utilisée pour mettre en évidence les termes qui sont le moins souvent employés, ainsi ne trouve-t-on presque jamais de termes liés à Dieu avec l'amour.

3) *Conclusions*

Ces différentes études apportent véritablement, si ce n'est une preuve, du moins, par les résultats obtenus, un argument très fort en faveur de la théorie dite cornélienne, et posent un vrai doute. Des ressemblances aussi troublantes entre des œuvres d'hommes si différents ne peuvent pas laisser insensible, même si l'on n'y croit pas au premier abord ; d'autant plus qu'on nous montre que les résultats lors des comparaisons avec d'autres dramaturges de l'époque ne sont pas du tout les mêmes...

Si on ne veut pas admettre la collaboration, on doit cependant reconnaître l'extrême influence de Corneille sur Molière, qui a été avant tout comédien et qui a joué énormément de pièces signées du nom de ce grand dramaturge. Molière aurait réussi à écrire presque comme son aîné, démentant ainsi l'affirmation de Racine quant à l'inimitabilité du style de Corneille.

Monsieur Dominique Labbé a bien sûr vérifié maintes et maintes fois ses résultats, et conclut d'ailleurs son livre *Corneille dans l'ombre de Molière* en soulignant que ses premières analyses étaient tellement probantes que l'on aurait pu se dispenser des deux autres. Il souligne également qu'il ne faut pas pour autant retirer tout mérite à Molière, qui a mis en scène et défendu des pièces qui, pour certaines, lui valurent beaucoup d'ennuis.



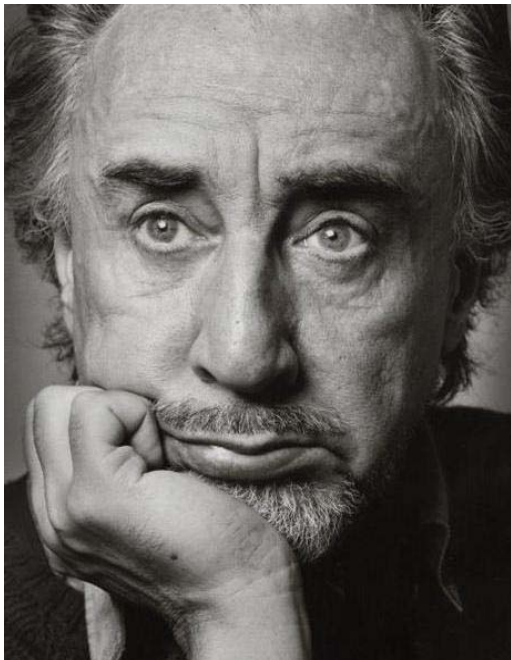
?

Montage d'un tableau représentant Molière écrivant, en fond, et d'un portrait de Corneille

ROMAIN GARY - EMILE AJAR

Comme nous l'avons dit précédemment, « deux » autres auteurs présentent les mêmes ressemblances entre leurs œuvres, d'après les résultats du même professeur, que pour l'étude comparative des pièces de Molière et de Corneille : il s'agit de Romain Gary et d'Emile Ajar.

1) L'affaire Gary-Ajar



L'histoire est ici totalement différente, puisque la solution est connue et claire : Emile Ajar est un des pseudonymes du romancier Romain Gary, né en 1914.

Au début des années soixante-dix, il est déjà un écrivain très connu : en effet, sa carrière, qui a commencé avec son premier roman *Education européenne*, paru en 1945, est couronnée de succès. Il a notamment reçu le prix Goncourt dès 1958 pour *Les Racines du ciel*.

L'écrivain ne fait alors plus le bonheur du grand public : en 1973, il a déjà écrit dix-neuf romans et il éprouve le sentiment de ne plus surprendre personne. C'est là que commence l'une des plus grandes supercheries, en tout cas prouvée, de l'histoire littéraire française. Romain Gary s'attelle à une autre de ses œuvres : *Gros câlin*, mais sa parution en 1974 est faite sous le nom d'Emile Ajar. Ce livre, considéré comme un premier roman, connaît un grand succès, mais très vite un doute se propage quant à la véritable identité de son auteur. Pour éviter tout soupçon,

Romain Gary donne vie à Emile Ajar ; et c'est son jeune cousin Paul Pavlovitch qui devra assumer ce rôle.

En novembre 1975, Emile Ajar reçoit pour *La Vie devant soi* le prix Goncourt, prix qu'un écrivain ne peut recevoir qu'une seule fois. Mais Paul Pavlovitch, bientôt identifié comme le neveu de Gary, affirme n'être pour rien dans la rédaction du texte, ce qui ne compromet pas le succès triomphal du roman, qui se vend à plus d'un million d'exemplaires. Le quatrième et dernier roman signé Emile

Ajar sera *L'Angoisse du roi Salomon* en 1979. Romain Gary, qui avait continué à écrire sous son vrai nom durant cette période, se suicide le 2 décembre 1980. Dans *Vie et mort d'Emile Ajar*, un texte écrit le 21 mars 1979, il avait lui-même annoncé pour conclure : « Je me suis bien amusé. Au revoir et merci. » Le 30 juin 1981, un communiqué de l'AFP dévoile la véritable identité d'Ajar.

2) *La Vie devant soi et Clair de femme*

Nous avons réalisé deux commentaires sur les incipits de ces deux romans, parus respectivement en 1975 et 1977 et publiés sous le nom d'Ajar pour le premier et de Gary pour le second, afin de voir quelles ressemblances au niveau stylistique il pouvait exister entre eux.

a) *La Vie devant soi, Momo et Rosa*



La rue de Belleville.

Nous allons étudier le plus célèbre des romans de Romain Gary (1914-1980), *La vie devant soi*. Ce livre, paru sous le nom d'Émile Ajar, reçoit le prix Goncourt en 1975. Il retrace une partie de la jeunesse de Mohammed, enfant arabe d'une dizaine d'années, surnommé Momo, qui vit dans une pension clandestine parisienne pour les enfants de prostituées, dirigée par Madame Rosa, une vieille Juive. Nous observons une véritable histoire d'amour entre ces deux personnages, le jeune Musulman soutiendra madame Rosa jusqu'à la fin de ses jours. Nous commenterons l'incipit du roman, le premier chapitre permettant de situer le lieu, à Paris, dans le quartier de Belleville, mais aussi les personnages, bien évidemment les deux protagonistes, Madame Rosa et Mohammed, mais aussi des personnages secondaires qui joueront un rôle dans la suite du récit. Enfin, nous étudierons le style très percutant utilisé par l'auteur.

Le narrateur, Momo, nous décrit madame Rosa tout d'abord physiquement puis nous présente ses origines. Même si son portrait moral n'est pas dressé, nous pouvons aisément en deviner certains traits qui seront confirmés par la suite du roman. Il insiste beaucoup sur le poids de la vieille Madame Rosa, le lecteur imagine donc dès la première phrase que cette femme est grosse «avec tous ces kilos qu'elle portait sur elle et seulement deux jambes», l'adverbe «seulement» indique un degré d'insuffisance et «tous ces kilos» traduit ainsi un fardeau. De plus, le narrateur nous indique que sa santé n'est pas bonne, sans doute due à sa très probable obésité. Le jeune garçon glisse avec une touche d'humour qu'elle «aurait mérité un ascenseur». Son nom Rosa apparaît ainsi comme un contraste à son physique disgracieux. C'est une Juive polonaise, ancienne prostituée, elle s'est défendue au Maroc et en Algérie et, de ce fait, a acquis une parfaite connaissance de la langue arabe. Elle a continué à se prostituer à son arrivée en France, mais confrontée à son âge et à sa croissante laideur, elle a dû arrêter et a finalement décidé d'ouvrir une pension pour «enfants de putes». Durant la seconde guerre mondiale, elle fut déportée à Auschwitz, une expérience qui lui laissera forcément des traces que l'on observera dans la suite du roman.

Elle a été en effet très touché par ce conflit : dès que l'on toque à la porte, elle ne peut s'empêcher de sursauter, se remémorant la fois où on l'avait emmenée au Vélodrome d'hiver, c'est

pourquoi elle s'est forgée une nouvelle identité à l'aide de faux-papiers ; autre fait étrange, elle conserve un portrait d'Hitler sous son lit.

Dans la suite du roman, madame Rosa s'amollit et meurt finalement, elle sera accompagnée et soutenue durant toute sa sénescence par le narrateur, Mohammed, qu'elle considère comme ce qu'elle a «de plus cher au monde».

Mohammed, surnommé Momo, est le narrateur de l'histoire; orphelin de cœur (il n'a jamais connu ses parents : sa mère, une prostituée, est morte et son père a été interné en psychiatrie pour avoir assassiné cette dernière) ; c'est l'enfant le plus âgé, le plus mature et le seul qui restera tout au long du roman à la pension. Ce jeune Musulman a été renvoyé de l'école et est très émotif, rien que dans ce premier chapitre, il pleure au moins quatre fois. Il a une vue plutôt pessimiste de la vie, « chienne de vie », c'est d'ailleurs pour cela qu'il ne veut pas se droguer, car selon lui, se droguer rend les gens heureux, or la vie est loin d'être aussi joyeuse. Il est très attaché à Rosa car il n'a qu'elle (il la connaît depuis l'âge de trois ans), et a une vue très naïve de la société, il s'est même fait convaincre qu'être juif est un fardeau «elle nous le rappelait à chaque fois qu'elle ne s'en plaignait pas d'autre part, car elle était également juive». Il incarne un des principaux thèmes des romans d'AJAR : le problème d'identité, il ne connaît pas son âge réel «je devrais avoir à peu près 7 ou 8 ans» et ne connaît pas ses parents, il fait d'ailleurs des choses stupides pour se faire remarquer espérant ainsi les faire venir, il feint par exemple d'avoir mal au ventre ou encore des convulsions, il en vient même à «chi[er] partout dans l'appartement» ou fait du vol à l'étalage en faisant exprès de se faire prendre. Il restera auprès de Madame Rosa malgré des propositions qui lui auraient permis une vie meilleure, il fera tout son possible pour réaliser le souhait de madame Rosa, qui ne veut pas aller à l'hôpital, au prix d'un gros mensonge (sa famille l'emmène en Israël) et va permettre à celle-ci de mourir dans son « trou juif ». Mais, il ne pourra s'occuper de cette femme, lors de sa défaillance, tout seul : de nombreux personnages vont l'aider, quelques-uns sont présentés dans ce premier chapitre.

Il s'en dégage trois principaux : Monsieur Hamil, Moïse et madame Lola. Le premier, marchand de tapis ambulant, avec lequel Mohammed tient une conversation, est un vieux (plus de soixante ans) Musulman «il portait une djellaba grise». Il paraît comme le seul adulte cultivé proche du narrateur : il a toujours sous la main *Les Misérables* de Victor Hugo qu'il connaît par cœur. C'est avec lui que Momo va apprendre et grâce à lui qu'il voudra devenir écrivain. Au contraire de Momo, il incarne la vision optimiste de la vie, une sorte de Pangloss des temps modernes, «Monsieur Hamil, pourquoi vous avez toujours le sourire? ». Au fil et à mesure du roman, il devient aveugle.

Le second est un enfant juif qui habite, du moins au début du roman, à la pension de Madame Rosa. Il sera ensuite adopté par une famille juive et aidera Momo à s'occuper de Madame Rosa lorsqu'elle sera très malade. C'est le seul ami du narrateur qui est nommé, mis à part Banania, le Noir qui souriait pour aucune raison.

Enfin, madame Lola est un travesti sénégalais, ancien champion de boxe, il se défend au Bois de Boulogne où elle a grand succès. Elle sera d'une grande aide lorsque madame Rosa sera sur le point de défaillir : en plus d'aider Momo à s'occuper d'elle, il leur donnera de l'argent pour leur permettre de manger et de survivre.

Ces trois personnes représentent la société vue par le jeune narrateur, ceci témoigne d'une diversité religieuse (à noter que les *Blancs* ne sont pas représentés dans ce chapitre mais le seront par la suite) et même professionnelle. Le narrateur rend un hommage à chacun des personnages qui l'ont marqué en affirmant son désir de les retrouver, l'auteur donne ainsi la sensation au lecteur d'une histoire vraie.

Cette œuvre d'Émile Ajar ne se distingue pas par son histoire, certes très originale et émouvante, mais par son écriture : le style est volontairement pas bon, en témoignent de mauvaises tournures comme la première phrase « La première chose que je peux vous dire », de nombreuses répétitions

présentes tout au long du texte... Mais voilà ce qui fait l'éclat de l'œuvre, c'est qu'un enfant, Mohammed, parle, accompagné de son langage familier « j'ai chié », « caca », « enfant de putes »,... ou encore de fautes de vocabulaire, « proxynète » à la place de proxénète ; de plus, ce narrateur surprenant s'adresse directement au lecteur avec la présence du pronom personnel vous « la première chose que je peux vous dire » (ce modalisateur est un bon exemple), ceci sensibilise le lecteur et permet de confronter la société à un nouveau point de vue, celui d'un enfant désigné comme perturbé et issu d'un quartier défavorisé. Voilà, la magie de ce texte est due au langage, ce qui correspond à ce qu'a dit récemment José Saramago, prix Nobel de littérature : « De la même manière que le corps se compose de 70% d'eau, la littérature doit comporter 70% de langage, les 30% restant sont l'histoire, les personnages ».

En ce qui concerne l'histoire, le chapitre 1 sert de scène d'exposition : en reprenant la théorie de Milan Kundera qui considère que le roman est une expérience où l'on observe l'évolution d'un individu dans un milieu précis, les conditions expérimentales sont en place : le lieu, Belleville, ainsi que les personnages principaux ont été présentés. Mais, en plus de cela, l'auteur nous expose déjà les principaux thèmes que l'on observera dans le roman : la mort, avec de nombreuses phrases contenant le verbe mourir « elle disait qu'un jour elle allait mourir dans l'escalier », « il me reste très peu de temps, je vais mourir avant »; la peur montrée ici par la réaction des enfants ou par le mot peur, « si je continuais c'était l'Assistance publique et là j'ai eu peur »; les religions et ainsi les questions de racisme « il y avait beaucoup d'autres Juifs, Arabes et Noirs » ; l'identité, sujet souvent présent dans les romans d'AJar, est dans ce roman exposée sur les questionnements de Momo au sujet de ses parents mais aussi sur son âge, il s'avérera, d'ailleurs, avoir quatre ans de plus que ce qu'il croyait. Dans ce chapitre, on observe de nombreuses approximations de la part de Momo à ce sujet « je devais avoir trois ans », « j'ai cessé d'ignorer à l'âge de trois ou quatre ans », « en ce moment je devais avoir sept ans ou peut-être huit » ; les prostituées qui encadrent la vie du narrateur « elle s'était défendue au Maroc et en Algérie », « Madame Lola était une travestie de quatrième étage qui travaillait au Bois de Boulogne » ; enfin le thème de l'amour dévoilé dans ce chapitre par l'attachement mutuel de Mohammed et de madame Rosa.



Affiche de l'adaptation cinématographique datant de 1977, réalisée par Moshé Mizrahi.

La Vie devant soi est donc raconté par le jeune Momo, mais est construit autour de Madame Rosa, sur ses hantises, sa détresse,... C'est de ces deux personnages que naît l'émotion, d'une part pour un pauvre gamin attachant et touchant par sa naïveté et son histoire et d'autre part pour cette vieille femme qui ne l'est pas moins, elle qui a sa vie derrière elle et sa mort au devant. Mais ce qui a contribué au succès du roman est avant tout l'écriture, et la vision du monde d'un enfant que la vie n'a

pas gâté, avec beaucoup de touches comiques qui contribuent au regard naïf sur le monde. Cette histoire montre qu'il y a de la place pour l'amour dans ce monde de fous.

b) Clair de femme, la rencontre

Nous allons maintenant étudier le début d'un autre roman de Romain Gary : *Clair de femme*. Ce livre, contrairement à *La vie devant soi* (1975), est paru sous le nom de Romain Gary en 1977. A travers ce roman, nous sommes les témoins d'une rencontre hasardeuse entre deux êtres au destin tragique : Michel, un commandant de bord dont la femme, Yannik, est en train de mourir, et Lydia, une femme mûre qui vient de perdre sa fille et son mari dans un accident de voiture. Mais cette rencontre se termine aussi brièvement qu'elle a commencé et le lecteur n'attend qu'une chose : que les deux amoureux d'un soir se revoient un jour. Le premier chapitre de ce roman

est le plus important car il décrit, de façon très précise, la première rencontre entre ces deux individus. Nous commenterons donc ce premier chapitre qui met en place à la fois le décor mais aussi la trame de l'histoire. Nous décrirons les personnages présents pour aboutir aux messages délivrés par l'auteur.

L'auteur nous présente trois personnages dans cet incipit. Tout d'abord, Michel, le narrateur : le premier adjectif qui le qualifie est « clownesque » et la première image qu'on a de lui est celle d'un homme maladroit. Commandant de bord chez Air France depuis 17 ans (carrière succédant à celle de barman), il était sur le point d'embarquer en direction de Caracas afin de se créer une nouvelle vie, la femme qui l'aimait, Yannik, étant mourante. Il est assez fort physiquement « épaules solides », « regard ancré » et garde un « air tranquille » malgré des événements troublants. Nous apprenons, plus tard, que cette tranquillité n'est qu'un masque qui cache une grande émotivité, à tel point qu'il en devient confus et oublie qu'il est en France, « je fus informé qu'on est en France, ici ». C'est un homme très aimable qui s'est fait beaucoup d'amis : « Après dix sept ans d'Air France, j'étais connu de la plupart des équipages, et je voulais éviter les rencontres et les questions amicales ».

Lors de sa rencontre avec Lydia, il se compare à « des pneus crevés qui peuvent faire encore mille kilomètres » : cette métaphore montre le coup de foudre qu'il a eu pour Lydia ; mais il reste quand même fidèle à Yannik et se sent mal à l'aise quand il se dirige vers l'appartement de Lydia : « il y avait des moments de panique et de vide, des échos de rires, des flambées de souvenirs, la fatigue fouillait dans le tout-à-l'égout de la mémoire et rejetait à la surface des bribes de bonheur. Le reste était angoisse et remords. Chaque minute qui passait paraissait arrachée à un temps fossilisé. ». Ce passage le montre bien : ces vieux souvenirs de Yannik ressurgissent soudainement et ce moment qui devait être agréable tourne au calvaire où chaque minute est un nouveau supplice. Il va même se considérer comme l'assassin de Yannik « qui revenait rôder sur les lieux du crime ».



Cette Lydia est une femme de 43 ans (N.B : même âge que Michel) qui vient de perdre son enfant et son mari quand elle rencontre Michel. Malgré un visage où « les rides se creusaient [...] et les années se posaient » et un air essoufflé « comme si elle était sorti d'un sauna », Lydia semble être en harmonie totale avec son âge : « un visage qui semblait avoir attendu les cheveux blancs pour réussir ce que la jeunesse et l'agrément des traits n'avaient fait qu'esquisser comme une promesse » ; malgré une tendance à se vieillir (elle se fait passer pour une femme de 47 ans). Le narrateur s'attarde sur ses longs cheveux blancs tombant « jusqu'aux épaules » et qu'il qualifie de « tumultueux », référence poétique qui nous montrent l'inclination qu'il a pour elle ; et sur son visage

aux pommettes « hautes » et « saillantes », aux yeux noirs « étrangement éloignés », « coulés en profondeur », « enfoncés dans leur royaume d'ombre » sorte de miroirs reflétant sa peine et sa douleur, et aux sourcils « très droits », « qui portent au milieu une ride profonde, comme un oiseau porte son corps », le mot ride profonde masquant sans doute le mot cicatrice douloureuse définit comme un lourd fardeau. A première vue, c'est une femme généreuse qui « vient au secours » de Michel et « le tire de la galère » ; mais cette générosité va se transformer en froideur à l'égard de ce dernier à la fin de ce premier chapitre : ceci est paradoxale pour une femme qui fut longtemps heureuse d'après son sourire. Tout objet lié à Lydia reflète son état psychologique actuel : son long manteau gris, son appartement terne, avec son « air de flûte indienne » qui « grignotait la solitude » et ses fauteuils de velours « gris-vert », montre, une fois encore, une personne triste. Mais tout espoir n'est pas perdu tel que le montre la clarté du salon et la présence de fleurs dans son appartement.

Le senior Galba, troisième personnage principal, est lui un vieil homme vêtu d'un « poil de chameau au col relevé » et d'un « vieux Borsalino au large bord rabattu gaillardement sur l'œil », aborde Michel « *Chez Arys* » le café où ce dernier et Lydia font plus ample connaissance. Il est persuadé d'avoir déjà vu Michel quelque part et ce dernier, pour s'en débarrasser, lui répond qu'ils se sont rencontrés à « Las Vegas en 1975 » et rajoute qu'il était « barman au *Sand's* », fait que Michel nie après à Lydia mais qui s'avère exact comme le prouve la phrase « les cheveux [...] avaient trente ans de plus que dans mes souvenirs ». C'est un « homme très élégant en poil de chameau et Borsalino » et qui se teint les cheveux pour « laisser du beau gris aux tempes ». Dresseur de chien, il dirige une troupe composée d'un assistant, d'un chimpanzé et de huit caniches, qui va de continent en continent pour présenter un « numéro mondialement connu », d'où son aptitude à parler sept langues différentes, dont le serbe, le russe et le polonais. Il n'a pour seul compagnon de route que son vieux caniche royal gris « taillé comme un Jardin de Le Nôtre », Mato Grosso, âgé de 14 ans (à peu près 70 ans pour un homme) et vestige d'une conquête passée. Ce chien, au museau gris et nu, l'accompagne depuis 13 ans mais n'est pas dressé pour le spectacle : c'est une sorte de lot de consolation que lui a laissé une de ses anciennes conquêtes. Sentant que son maître est âgé, il est anxieux et a peur de se retrouver seul après, comme son maître (tel chien tel maître). Le narrateur s'attarde une nouvelle fois sur le visage d'un autre personnage : il est fasciné par ce nez « conséquent » et « méphistophélique » aux « superbes narines » qui est en avant « comme pour chercher des secours » et ses « yeux d'olives noires » qui permettaient de lire « une vieille de Casanova » et montraient un regard anxieux. Ce Casanova ne fut pas épargné par la solitude malgré de nombreuses conquêtes : il estime que « trop c'est personne ». Suite à de nombreux ennuis avec les assurances, son « œuvre d'une vie » n'a pas suffi à l'enrichir et l'a progressivement ruiné : ces événements furent pour lui « trois petits infarctus » qui l'ont mené à sa mort « smrt » (N.B : mort en serbe) c'est-à-dire à sa perte.

Ces trois personnages, qui représentent différents domaines professionnels et sociaux, sont en réalité très liés à cause de leurs vies brisées. Dans ce premier chapitre, le sujet récurrent est la perte d'un être cher : en effet chaque personnage a perdu (Lydia), est en train de perdre (Michel), ou perdra (Galba) une personne comptant énormément pour lui.

Ce premier chapitre est une vitrine du style de Romain Gary : remarquons d'abord que c'est un des personnages qui raconte l'histoire et que chaque personnage est différent en raison de son statut professionnel et social. Nous pouvons ensuite constater que ce premier chapitre est construit sur un système tertiaire : trois personnages présentés (Michel, Lydia, le senior Galba), trois personnages en arrière plan (le chauffeur de taxi, le serveur et l'adolescent près du juke-box), trois lieux différents (l'aéroport, le café et l'immeuble de Lydia), trois vies brisées, trois infarctus, trois traductions slaves pour le mot mort, trois cartes de visites...

Gary utilise un style familier tel que le montrent des mots ou des expressions « qu'est-ce que tu fous ici ? », « prendre un Bic », « un calva, une fine », « salaud » ou des mots empruntés aux langues étrangères : « cheers », « senor », « smrt ». De plus l'auteur fait ressortir les états psychologiques actuels de ses personnages à la fois dans des descriptions physiques centrées sur le buste des personnages et aussi dans des choses que côtoient régulièrement les personnages tel que l'appartement de Lydia, le chien angoissé de Galba, le temps (ici pluvieux) sous lequel se déroule l'action ; ce premier chapitre est très grisé, triste, malheureux mais nous pouvons distinguer au loin un rayon d'espoir nais-

sant de la rencontre entre Michel et Lydia. A travers ce roman nous pouvons déjà anticiper et proposer une liste probable de sujets récurrents : la vie, la mort, le couple et la séparation. Notons pour finir l'hommage, certes très discret mais présent, de l'auteur à ses origines slaves : le senior Galba parle le serbe, le russe et le polonais et trouve que « ce sont les Slaves qui ont su trouver le meilleur nom, le son le plus vrai, pour la chose » (ici la mort).

Clair de femme est un chant d'amour qui montre l'importance de cette troisième dimension, de ce lien indestructible entre l'homme et la femme qu'est le couple. A travers les yeux de Michel, nous pouvons observer la débâcle d'un homme face à la disparition de sa femme et le désir de trouver puis retrouver l'amour d'une autre.

c) Conclusion

Ces deux textes n'ont pas le même style d'écriture, le même langage. Le premier est celui d'un enfant issu d'un milieu modeste, ayant reçu une éducation succincte. Le deuxième est celui d'un homme adulte. Cependant, au-delà du style, on retrouve des thèmes similaires et universels, principalement l'amour. L'amour entre deux personnes qui, *a priori*, ne devraient pas se lier de cette façon, du fait de leur origine, de leur âge ou de la situation dans laquelle elles se trouvent. On retrouve aussi celui de la mort, très développé dans ces deux textes, mais également dans toute l'œuvre de Gary, qui d'ailleurs se suicidera, ce qui dénote un très certain mal-être. Enfin, ces deux extraits nous montrent pareillement un attachement de l'auteur à ses origines à la fois juives et slaves.

3) L'expérience d'attribution d'auteur

Nous avons demandé à Dominique Labbé, le 14 janvier, s'il lui était possible de nous transmettre les logiciels permettant de réaliser les calculs de la distance intertextuelle que nous souhaitions utiliser pour comparer le début des œuvres¹ d'AJar et Gary contemporaines ainsi que celui de *La Promesse de l'aube*, datant de 1960 et donc antérieur à la période « ajarienne ».

Cependant, à la réception du CD-ROM, un mois après l'envoi depuis Montréal le 17 janvier, il nous a été impossible de faire marcher correctement les logiciels qu'il contenait car ils étaient dépourvus d'interface et parfois impossibles à mettre en marche. Nous avons donc demandé à Monsieur Labbé s'il pouvait faire pour nous les calculs.

Le corpus est donc constitué des débuts (24 000 premiers mots) des 4 romans parus sous le nom d'AJar (*Gros Câlin*, *Pseudo*, *L'angoisse du roi Salomon* et *La Vie devant soi*) et 6 autres sous le nom de Gary (*La Promesse de l'aube*, *Charge d'âme*, *Clair de femme*, *Chien blanc*, *Les Cerfs-volants* et *Au-delà de cette limite...*) chacun divisé en deux parties égales.

¹ Il est à noter que, les textes n'étant pas encore dans le domaine public, D. Labbé a dû numériser lui-même les textes à comparer page après page et nous les a transmis.

Voici donc le tableau que l'on peut dresser premièrement, après avoir usé de la lemmatisation sur les textes.

	Date	Vocables
Ajar :		
Gros Calin1	1974	1 813
Gros Calin2	"	1 854
Pseudo1	1976	1 935
Pseudo2	"	1 971
Roi Salomon1	1979	1 649
Roi Salomon2	"	1 568
Vie devant soi1	1975	1 331
Vie devant soi2	"	1 310
Gary :		
Charge d'âme1	1977	2 449
Charge d'âme2	"	2 387
Cerfs volants1	1980	2 120
Cerfs volants2	"	2 159
Chien blanc1	1970	2 298
Chien blanc2	"	2 315
Clair de femme1	1977	1 847
Clair de femme2	"	1 874
Promesse de l'aube1	1960	2 265
Promesse de l'aube2	"	2 360
Au-delà de cette limite1	1974	2 175
Au-delà de cette limite2	"	2 042

On remarque premièrement que le nombre de vocables différents chez Ajar est systématiquement inférieur à celui de Gary, sauf pour *Clair de femme*. Il faut savoir que la densité du vocabulaire à l'oral est toujours inférieure à celle de l'écrit chez une même personne, et compte tenu des résultats des études vues précédemment, celui-ci semble justifié.

1) On passe ensuite les textes par les logiciels de calcul de la distance intertextuelle et l'on obtient, pour le corpus Ajar, les distances internes suivantes :

	GrosCâlin1	GrosCâlin2	Pseudo1	Pseudo2	Salomon1	Salomon2	VieDevant1	VieDevant2
GrosCâlin1	0,000	0,211	0,268	0,266	0,290	0,298	0,301	0,315
GrosCâlin2	0,211	0,000	0,258	0,252	0,275	0,292	0,290	0,304
Pseudo1	0,268	0,258	0,000	0,221	0,296	0,304	0,300	0,298
Pseudo2	0,266	0,252	0,221	0,000	0,284	0,295	0,287	0,287
Salomon1	0,290	0,275	0,296	0,284	0,000	0,223	0,237	0,255
Salomon2	0,298	0,292	0,304	0,295	0,223	0,000	0,233	0,242
VieDevant1	0,301	0,290	0,300	0,287	0,237	0,233	0,000	0,186
VieDevant2	0,315	0,304	0,298	0,287	0,255	0,242	0,186	0,000

On peut ici remarquer une assez bonne homogénéité, la distance moyenne étant de 0.27. De plus, tous les extraits d'un même livre sont séparés d'une distance inférieure à 0.25. On remarque également une parenté nette entre *La Vie devant soi* et *L'Angoisse du roi Salomon* : on peut ainsi conclure que Gary a créé et reproduit un style à 4 reprises sur 5 années.

2) Les distances dans le corpus Gary :

	Ame1	Ame2	Cerfs1	Cerfs2	Chien 1	Chien 2	Clair 1	Clair 2	Promesse1	Promesse2	Au-dela1	Au-dela2
Ame1	0,000	0,267	0,336	0,328	0,361	0,344	0,387	0,377	0,345	0,358	0,346	0,350
Ame2	0,267	0,000	0,331	0,322	0,341	0,321	0,358	0,349	0,341	0,360	0,333	0,339
Cerfs1	0,336	0,331	0,000	0,234	0,296	0,330	0,316	0,304	0,276	0,285	0,277	0,274
Cerfs2	0,328	0,322	0,234	0,000	0,297	0,320	0,307	0,302	0,283	0,297	0,279	0,269
Chien1	0,361	0,341	0,296	0,297	0,000	0,291	0,310	0,306	0,305	0,322	0,286	0,285
Chien2	0,344	0,321	0,330	0,320	0,291	0,000	0,357	0,350	0,328	0,343	0,318	0,332
Clair1	0,387	0,358	0,316	0,307	0,310	0,357	0,000	0,214	0,339	0,353	0,285	0,260
Clair2	0,377	0,349	0,304	0,302	0,306	0,350	0,214	0,000	0,334	0,351	0,281	0,263
Promesse1	0,345	0,341	0,276	0,283	0,305	0,328	0,339	0,334	0,000	0,233	0,296	0,301
Promesse2	0,358	0,360	0,285	0,297	0,322	0,343	0,353	0,351	0,233	0,000	0,314	0,302
Au-dela1	0,346	0,333	0,277	0,279	0,286	0,318	0,285	0,281	0,296	0,314	0,000	0,246
Au-dela2	0,350	0,339	0,274	0,269	0,285	0,332	0,260	0,263	0,301	0,302	0,246	0,000
Moyenne	0,345	0,333	0,296	0,294	0,309	0,330	0,317	0,312	0,307	0,320	0,296	0,293

Ce corpus, au contraire, est beaucoup plus hétérogène : la distance moyenne y est de 0.31. Toutes les distances y sont supérieures à 0.25 si l'on excepte les distances entre extraits de même œuvre et quelques distances sont supérieures au seuil de 0.35, à partir duquel l'hypothèse de deux auteurs différents devient de plus en plus envisageable.

Gary a utilisé plusieurs fois l'image du caméléon pour se qualifier lui-même, son œuvre « officielle » en apporterait ainsi un témoignage...

3) Les distances entre les corpus Ajar et Gary :

	GrosCâlin1	GrosCâlin2	Pseudo1	Pseudo2	Salomon1	Salomon2	VieDevant1	VieDevant2
Ame1	0,375	0,372	0,397	0,384	0,405	0,427	0,439	0,457
Ame2	0,354	0,351	0,377	0,357	0,384	0,394	0,413	0,424
Cerfs1	0,301	0,300	0,333	0,316	0,319	0,348	0,357	0,374
Cerfs2	0,302	0,298	0,328	0,312	0,317	0,345	0,354	0,370
Chien1	0,286	0,292	0,321	0,310	0,334	0,345	0,355	0,363
Chien2	0,335	0,333	0,349	0,349	0,373	0,391	0,394	0,412
Clair1	0,269	0,263	0,293	0,275	0,290	0,288	0,307	0,309
Clair2	0,272	0,266	0,286	0,266	0,289	0,285	0,303	0,309
Promesse1	0,329	0,330	0,360	0,344	0,359	0,380	0,384	0,406
Promesse2	0,341	0,339	0,371	0,359	0,368	0,388	0,396	0,416
Au-dela1	0,281	0,278	0,300	0,289	0,312	0,322	0,340	0,348
Au-dela2	0,284	0,280	0,293	0,280	0,309	0,322	0,334	0,345
Moyenne	0,311	0,309	0,334	0,320	0,338	0,353	0,365	0,378

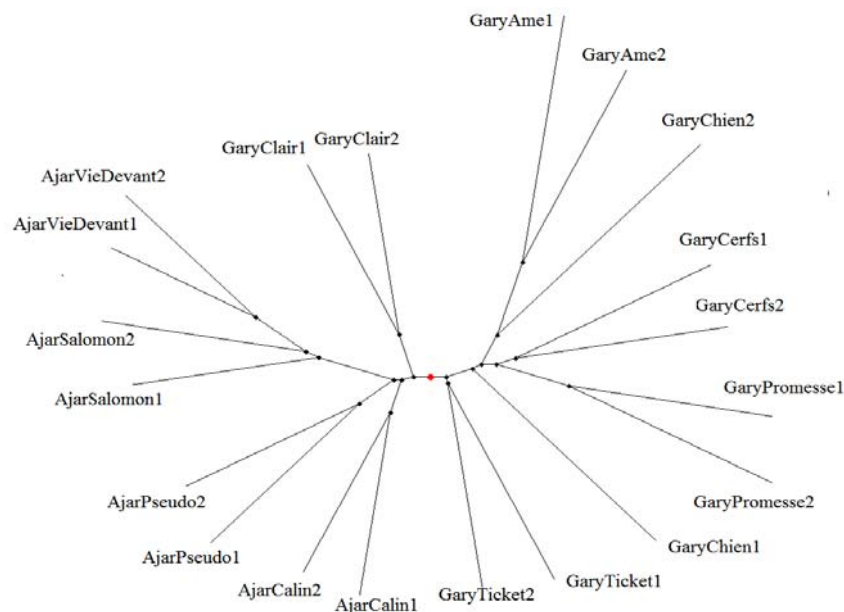
Ici, toutes les distances sont élevées, certaines même supérieures à 0.4.

La Vie devant soi est manifestement le plus éloigné (cf. ligne des moyennes). Il semble donc compréhensible que les jurés du prix Goncourt en ayant lu le début de cette œuvre n'aient pas remarqué la supercherie, même en connaissant bien l'œuvre de Gary.

Clair de femme est le Gary le plus proche de chaque Ajar, et il est, de même qu'*Au-delà de cette limite*, spécialement proche de *Gros-Câlin* et de *Pseudo*.

4) La classification arborée :

Il s'agit d'une des très nombreuses méthodes de classifications et est le plus souvent utilisée en génétique : elle repose sur les distances euclidiennes, qui permettent de dresser un « arbre » qui représente les positions respectives des textes les uns par rapport aux autres, et ensuite les meilleurs groupements possibles entre eux. Cette méthode a été mise au point par un certain Luong.



Ici, chaque texte est représenté par une "feuille" terminale. Les points noirs représentent les meilleurs groupements, c'est-à-dire lorsque les distances entre les membres sont les plus faibles et les distances avec les autres sont les plus grandes.

Le point rouge, qui se trouve au milieu du segment reliant les parties gauche et droite de l'arbre est appelé « racine ».

On remarque qu'à part pour *Chien blanc*, les extraits d'un même ouvrage sont tous les deux mariés directement l'un avec l'autre.

Le point rouge sépare, lui, deux parties dans lesquelles on retrouve, d'une part, les œuvres d'Ajar, et d'autre part celles de Gary. En revanche on voit que *Clair de femme* se retrouve dans la première et qu'*Au-delà de cette limite* en est très proche : cette classification résume donc les conclusions que l'on a pu tirer des tableaux précédents.

On peut donc dire que *Clair de femme* est plus de l'Ajar que du Gary, et remarquer également que *La Promesse de l'aube*, roman autobiographique de 1960, est rattachée aux *Cerfs-volants*, également autobiographique mais de 20 ans plus récent : on peut supposer qu'à la fin de sa vie, Romain Gary a souhaité revenir sur sa jeunesse et la guerre, retrouvant ainsi son style d'antan.

Remarque : Dominique Labbé a réalisé les mêmes expériences, en bien sûr plus approfondi, sur le même corpus (excepté *La Promesse de l'aube*) mais sur les textes complets : tous les résultats se trouvent à l'adresse suivante :

<http://hal.archives-ouvertes.fr/docs/00/27/96/63/PDF/LabbeGaryAjar2008.pdf>

Vous trouverez les tableaux, équivalant à ceux présentés ici, qui en résultent en annexe 5, page 39.

5) Remarques conclusives sur cette étude :

- En utilisant seulement les premières pages de tous les romans, l'expérience montre que l'on n'aurait pas pu tirer la conclusion que Gary et Ajar ne font qu'un : il faut donc que de telles études statistiques soient opérées sur des œuvres entières si l'on souhaite avoir des résultats, d'une part réels, et pour ce cas, d'autre part, véritablement conclusifs.

- La ressemblance de deux romans de Gary avec les textes d'Ajar aurait pu mener à se demander s'il s'agissait d'un jeune auteur qui influençait son aîné, ou s'il s'agissait bien d'un même auteur. Car, étrangement, Ajar reste très stable et c'est Gary qui change beaucoup et qui se rapproche du premier, comme s'il n'avait pas réussi à maintenir une cloison étanche entre les deux styles. On peut donc en conclure que la distance intertextuelle ne fonctionne pas comme les empreintes digitales.

Dominique Labbé tient à rappeler que « la méthode ne prétend pas résoudre toutes les énigmes, mais qu'elle peut attribuer à coup sûr certains textes, quitte à laisser des cas non résolus ».

CONCLUSION

Une étude statistique, appuyée sur une étude biographique solide et sans zone d'ombre, pourrait parfaitement permettre de définir l'auteur d'un texte, si celui-ci est cependant comparé à d'autres datant de la même époque, préférentiellement dans un même répertoire littéraire fourni et avec un style de langue semblable.

Il faut donc que la thèse cornélienne soit reconnue comme plausible dans l'état actuel des connaissances sur Molière, et non pas la juger comme étant le fruit d'esprits dérangés. Cette affaire aurait bien évidemment un impact sur le plan littéraire mais aussi sur le plan politique : Molière est considéré comme un symbole républicain, on parle la langue de Molière et non celle de Corneille ! C'est probablement grâce à la République que Molière est considéré ainsi aujourd'hui.

Nous pourrions utiliser cette méthode d'analyse statistique pour d'autres affaires comme celle concernant madame de La Fayette et La Rochefoucauld, ce dernier étant suspecté d'être le véritable auteur de *La Princesse de Clèves* parue en 1678.

Notre étude avait également pour objectif d'illustrer la contribution mutuelle et l'enrichissement croisé de deux matières aussi éloignées l'une de l'autre que la littérature et les mathématiques.

Nous avons pu observer que les mathématiques statistiques, la logique et les programmes informatiques pouvaient contribuer, si ce n'est à résoudre, du moins à considérablement éclairer des problèmes littéraires. Réciproquement, les études approfondies nécessaires pour appréhender et modéliser la complexité d'une langue et de son expression littéraire sont, très certainement, des vecteurs de progrès des sciences mathématiques.

Espérons que toutes les intelligences mobilisées à la résolution, par des voies mathématiques, de ces grandes énigmes littéraires, enrichiront encore longtemps le savoir humain.

LEXIQUE

‘Federalist papers’ : il s’agit d’un recueil d’articles, écrit par James Madison, Alexander Hamilton et John Jay, publié en vue d’une promotion de la nouvelle Constitution des États-Unis d’Amérique. Il est paru dans les années 1787-1788 dans plusieurs États d’Amérique et contenait 85 articles. Ces articles soulignaient comment ce nouveau gouvernement allait opérer et de quelle manière ce type de gouvernement était le meilleur choix pour les États-Unis d’Amérique. (wikipédia)

Dix-septémistes : ‘experts’ de l’histoire du XVII^{ème} siècle, ils rejettent pour la plupart la thèse de Pierre Louÿs.

Moliéristes : défenseurs et connaisseurs de Molière, par opposition aux cornéliens.

Cornéliens : nom que se donnent et que l’on donne aux défenseurs de la thèse selon laquelle Pierre Corneille aurait écrit une partie de l’œuvre théâtrale attribuée à Molière. Ils sont également appelés ‘louÿsiens’, entre autres par François Rey, car ils reprennent l’idée initiale de P. Louÿs.

Vocabulaire : à l’origine qualifiant un terme signifiant quelque chose de précis, il symbolise ici un mot réduit à une forme unique.

ANNEXES

ANNEXE 1 : Réponse (extrait) de Denis Boissier suite à notre question au sujet de la datation de la pièce *L'Etourdi* de Molière.

Question : Pourriez-vous avoir la gentillesse de nous accorder un peu de votre temps pour répondre à la simple question suivante, au sujet de la pièce *L'Etourdi*. Vous indiquez en effet sur votre site internet "version parisienne 1658", pourtant l'on trouve sur toutes les autres sources la date de 1655, avec une représentation à Lyon ; nous aurions donc besoin de savoir si ces deux versions étaient différentes, et si ceci est prouvé.

Réponse :

« Pour les moliéristes, *L'Etourdi* et *Le Dépit amoureux* datent d'avant 1658 car ces comédies ont été jouées en province, respectivement à Lyon en 1655 et à Béziers en 1656. C'est vrai. Mais il faut s'entendre sur les mots. Certes, quelque chose a été joué sous ces titres, mais quoi ? Nous n'en savons rien. Au XVII^{ème} siècle une comédie pouvait évoluer beaucoup tout en conservant le même titre. On ajoutait des scènes, on en retirait, on changeait un personnage ou un dénouement... bref, une comédie qui n'était pas publiée pouvait se présenter successivement ou simultanément sous diverses formes. Or, nous ne connaissons pas le texte de ces deux comédies avant qu'elles n'aient été publiées après 1659.

« De plus, en 1658, la troupe de Madeleine Béjart et de Molière a passé six mois à Rouen, auprès des frères Corneille. Nous savons que les deux frères étaient de bons "arrangeurs" et nous savons aussi que la comédie des *Précieuses ridicules* n'est pas étrangère au fait que les Corneille lisaient à l'époque du séjour de la troupe à Rouen le roman de leur ami l'abbé de Pure *La Prétieuse* (oui, avec un "t"). Je ne m'éloigne pas de mon sujet en vous parlant des *Précieuses ridicules* car cette comédie présente la même difficulté de datation que *L'Etourdi* ou *Le Dépit amoureux*. *Les Précieuses ridicules* ont été jouées en 1659, mais plusieurs témoignages montrent qu'elles ont été jouées en province avant cette date, et nous savons aussi que la version provinciale était différente de la version parisienne (une spectatrice nommée Mlle Des Jardins a raconté la pièce provinciale et, à l'évidence, des changements ont été apportés pour la création parisienne). Les moliéristes datent *Les Précieuses ridicules* de leur représentation parisienne (1659) ; *L'Etourdi* et *Le Dépit amoureux* de leur représentation provinciale (1655 et 1656). Cette différence de traitement est illogique. Puisque nous ne savons rien des textes de 1655 et 1656 pourquoi en tenir compte et les prendre pour référence ? Ne serait-ce pas afin d'éviter toute interférence avec Corneille ?

« Nous pensons que, comme pour *Les Précieuses ridicules*, la datation de *L'Etourdi* et du *Dépit amoureux* doit être celle de leur création à Paris, ne serait-ce que parce que nous en possédons le texte.

Constatons aussi que La Grange écrit dans son *Registre*, à propos de *L'Etourdi* : « Comédie représentée pour la première fois à Paris sur le Théâtre du Petit Bourbon au mois de novembre 1658 par la troupe de Monsieur, frère du Roi », ce qui ne peut se justifier que si elle a été *réécrite*. C'était déjà l'avis du moliériste Victor Fournel pour qui *L'Etourdi* parisien n'est pas celui joué en province : « on ignore jusqu'à quel point Molière a pu modifier sa comédie dans l'intervalle. » (*Les contemporains de Molière, recueil de comédies rares ou peu connues*, 1863-1875, T. I, p. 5).

« En 1655 Pierre Corneille séjourne à Bourbon-Lancy, près de Lyon. La troupe a donc pu le fréquenter... et l'on conçoit que Corneille ait pu par amitié ou, déjà, par amour pour Marquise du Parc (elle jouera Hippolyte dans *L'Etourdi* version parisienne) améliorer la versification de cette comédie qui, ne l'oublions pas, est un démarquage d'une pièce italienne. Le vers 538 de *L'Etourdi* – « On n'a point pour la mort de dispense de Rome » – se trouve déjà presque mot pour mot dans l'un des ouvrages de Thomas à Kempis (au chapitre V de *Vallis liliorum*) paru en latin en 1574 ; or Thomas à Kempis est l'auteur de *L'Imitation de Jésus-Christ* que traduit alors Pierre Corneille. Relisez le monologue de Mascarille (acte V, sc. 1), l'entrevue des deux vieillards (acte III, sc. 4) et vous verrez que ces scènes sont d'une tout autre facture que le reste de la pièce. Voyez aussi ces vers de la scène 3 de l'acte IV :

[...] Quand on ne prend en dot que la seule beauté
 Le remords est bien près de la solennité.
 Et la plus belle femme a très peu de défense
 Contre cette tiédeur qui suit la jouissance. [...] (v.1473-1476)

« Pour Georges Couton, « Molière imite ici, consciemment semble-t-il, une tirade de *Mélie*, de Corneille. » (*Molière, oeuvres complètes*, Bibliothèque de la Pléiade, 1971, T. I, p. 1200). Tirade que voici :

[...] La beauté, les attraits, l'esprit, la bonne mine,
 Echaudent bien le cœur, mais non pas la cuisine,
 Et l'hymen qui succède à ces folles amours,
 Après quelques douceurs, a bien de mauvais jours. [...] (v. 117-120)

« On peut en dire autant du *Dépit amoureux*. Pour le moliériste René Jasinski, « une influence intervient aussi, dont on n'a pas assez tenu compte : un certain ton qui rappelle Corneille et plus particulièrement la comédie cornélienne. On reconnaît la propension aux généralités sentencieuses, la prédilection pour les cas de conscience plus ou moins subtils et prêtant aux psychologies raffinées, la recherche des effets de surprise dans un romanesque par moments poussé à l'extrême et qui se veut hors du commun : scènes plus soutenues qui rejoignent sinon la tragédie, du moins la comédie héroïque. Et c'est bien un des genres auxquels Molière aspire alors, auquel il s'essaiera dans *Dom Garcie*. L'emprise de Corneille sur sa formation en quelque sorte technique va donc plus loin qu'on ne l'a dit... » (*Molière*, 1969, p. 39). »

ANNEXE 2 : Jean Racine, *Eloge de Corneille* (1685) à l'Académie française. Il s'adresse à Thomas Corneille et vante les mérites de son frère.

« Il n'est pas besoin de dire ici combien l'Académie a... regardé la mort de Monsieur de Corneille comme un des plus rudes coups qui la pût frapper. Car bien que, depuis un an, une longue maladie nous eût privés de sa présence, et que nous eussions perdu en quelque sorte l'espérance de la revoir jamais dans nos assemblées, toutefois il vivait, et l'Académie, dont il était le doyen, avait au moins la consolation de voir dans la liste où sont les noms de tous ceux qui la composent, de voir, dis-je, immédiatement au-dessus du nom sacré de son auguste protecteur, le fameux nom de Corneille.

Et qui d'entre nous ne s'applaudissait en lui-même, et ne ressentait pas un secret plaisir d'avoir pour confrère un homme de ce mérite ? Vous, Monsieur, qui non seulement étiez son frère, mais qui avez couru longtemps une même carrière avec lui, vous savez les obligations que lui a notre poésie ; vous savez en quel état se trouvait la scène française lorsqu'il commença à travailler. Quel désordre ! quelle irrégularité ! Nul goût, nulle connaissance des véritables beautés du théâtre : les auteurs aussi ignorants que les spectateurs ; la plupart des sujets extravagants et dénués de vraisemblance ; point de mœurs, point de caractères ; la diction encore plus vicieuse que l'action, et dont les pointes et de misérables jeux de mots faisaient le principal ornement : en un mot toutes les règles de l'art, celles même de l'honnêteté et de la bienséance partout violées.

« Dans cette enfance, ou, pour mieux dire, dans ce chaos du poème dramatique parmi nous, votre illustre frère, après avoir quelque temps cherché le bon chemin et lutté, si j'ose ainsi dire, contre le mauvais goût de son siècle, enfin inspiré d'un génie extraordinaire, et aidé de la lecture des anciens, fit voir sur la scène la raison, mais la raison accompagnée de toute la pompe, de tous les ornements dont notre langue est capable, accorda heureusement le vraisemblable et le merveilleux, et laissa bien loin derrière lui tout ce qu'il avait de rivaux, dont la plupart, désespérant de l'atteindre, et n'osant plus entreprendre de lui disputer le prix, se bornèrent à combattre la voix publique déclarée pour lui, et essayèrent en vain par leurs discours et par leurs frivoles critiques de rabaisser un mérite qu'ils ne pouvaient égaler. La scène retentit encore des acclamations qu'excitèrent à leur naissance le Cid, Horace, Cinna, Pompée, tous ces chefs-d'oeuvre représentés depuis sur tant de théâtres, traduits en tant de langues, et qui vivront à jamais dans la bouche des hommes. A dire le vrai, où trouvera-t-on un poète qui ait possédé à la fois tant de grands talents, tant d'excellentes parties : l'art, la force, le jugement, l'esprit ? Quelle noblesse, quelle économie dans les sujets ! Quelle véhémence dans les passions ! Quelle gravité dans les sentiments ! Quelle dignité, et en même temps quelle prodigieuse variété dans les caractères ! Combien de rois, de princes, de héros de toutes nations nous a-t-il présentés, toujours tels qu'ils doivent être, toujours uniformes avec eux-mêmes, et jamais ne se ressemblant les uns aux autres ! Parmi tout cela, une magnificence d'expression proportionnée aux maîtres du monde qu'il fait souvent parler, capable néanmoins de s'abaisser, quand il veut, et de descendre jusqu'aux plus simples naïvetés du comique, où il est encore inimitable. Enfin, ce qui lui est surtout particulier, une certaine force, une certaine élévation qui surprend, qui enlève, et qui rend jusqu'à ses défauts, si on lui en peut reprocher quelques-uns, plus estimables que les vertus des autres. Personnage véritablement né pour la gloire de son pays ; comparable, je ne dis pas à tout ce que l'ancienne Rome a eu d'excellents tragiques, puisqu'elle confesse elle-même qu'en ce genre elle n'a pas été fort heureuse, mais aux Eschyles, aux Sophocles, aux Euripides, dont la fameuse Athènes ne s'honore pas moins que des Thémistocles, des Périclès, des Alcibiades, qui vivaient en même temps qu'eux.

« Oui, Monsieur, que l'ignorance rabaisse tant qu'elle voudra l'éloquence et la poésie, et traite les habiles écrivains de gens inutiles dans les Etats, nous ne craignons point de le dire à l'avantage des lettres et de ce corps fameux dont vous faites maintenant partie : du moment que des esprits sublimes, passant de bien loin les bornes communes, se distinguent, s'immortalisent par des chefs-d'oeuvre, comme ceux de Monsieur votre frère, quelque étrange inégalité que durant leur vie la fortune mette entre eux et les plus grands héros, après leur mort cette différence cesse. La postérité, qui se plaît, qui s'instruit dans les ouvrages qu'ils lui ont laissés, ne fait point de difficulté de les égaler à tout ce qu'il y a de plus considérable parmi les hommes, fait marcher de pair

l'excellent poète et le grand capitaine. Le même siècle qui se glorifie aujourd'hui d'avoir produit Auguste, ne se glorifie guère moins d'avoir produit Horace et Virgile. Ainsi, lorsque dans les âges suivants on parlera avec étonnement des victoires prodigieuses et de toutes les grandes choses qui rendront notre siècle l'admiration de tous les siècles à venir, Corneille, n'en doutons point, Corneille tiendra sa place parmi toutes ces merveilles. La France se souviendra avec plaisir, que sous le règne du plus grand de ses rois a fleuri le plus célèbre de ses poètes. On croira même ajouter quelque chose à la gloire de notre auguste monarque lorsqu'on dira qu'il a estimé, qu'il a honoré de ses bienfaits cet excellent génie ; que même deux jours avant sa mort, et lorsqu'il ne lui restait plus qu'un rayon de connaissance, il lui envoya encore des marques de sa libéralité ; et qu'enfin les dernières paroles de Corneille ont été des remerciements pour Louis le Grand.

« Voilà, Monsieur, comme la postérité parlera de votre illustre frère ; voilà une partie des excellentes qualités qui l'ont fait connaître à toute l'Europe. Il en avait d'autres, qui bien que moins éclatantes aux yeux du public, n'en sont peut-être pas moins dignes de nos louanges, je veux dire, homme de probité, de piété, bon père de famille, bon parent, bon ami ; vous le savez, vous qui avez toujours été uni à lui d'une amitié qu'aucun intérêt, non pas même aucune émulation pour la gloire n'a pu altérer. Mais ce qui nous touche de plus près, c'est qu'il était encore un très bon académicien. Il aimait, il cultivait nos exercices. Il y apportait surtout cet esprit de douceur, d'égalité, de déférence même, si nécessaire pour entretenir l'union dans les Compagnies. L'a-t-on jamais vu se préférer à aucun de ses confrères ? L'a-t-on jamais vu vouloir tirer ici aucun avantage des applaudissements qu'il recevait dans le public ? Au contraire, après avoir paru en maître, et, pour ainsi dire, régné sur la scène, il venait, disciple docile, chercher à s'instruire dans nos assemblées, laissait, pour me servir de ses termes, laissait ses lauriers à la porte de l'Académie, toujours prêt à soumettre son opinion à l'avis d'autrui, et de tous tant que nous sommes, le plus modeste à parler, à prononcer, je dis même sur des matières de poésie. »

ANNEXE 3 : Entrevue avec Monsieur François Rey

François Rey a commencé l'entretien en nous exposant ce qu'il savait et ce qu'il pensait de l'évolution de l'affaire : elle commence pour lui avec Pierre Louÿs - il met d'ailleurs au défi qui-conque de trouver quelqu'un ayant avant lui affirmé que Molière n'était pas l'auteur de ses pièces - qui, lorsqu'il a lancé son hypothèse (en pleine Première Guerre Mondiale), était sur la fin de sa vie, seul, malade et presque sans le sou. D'après Monsieur Rey, ce poète et écrivain reconnu était un grand connaisseur de Corneille, et n'appréciait pas spécialement Molière, comme en témoignent les correspondances qu'il a tenues avec Paul Valéry ou Lachèvre ; il a développé sa thèse sur des bases poétiques, croyant reconnaître le style de Corneille lors d'une lecture d'*Amphytrion*, et l'a enrichie à l'aide de faits historiques, comme on les retrouve dans la thèse actuelle.

Nous l'avons ensuite questionné à propos de l'éducation de Molière : il ne fait pour lui aucun doute que Molière n'est pas allé au collège de Clermont, ni qu'il ait suivi des études de droit ; mais le fait de ne pas avoir fait des études très poussées n'est pas un argument en soi. Il avance ensuite que Molière a fréquenté dans sa jeunesse ce que l'on appelle les petites écoles, où l'on recevait des cours permettant d'avoir un niveau comparable à celui des anciens certificats d'études. Il affirme cependant que Molière avait une connaissance assez poussée en latin, en quelque sorte l'anglais de l'époque, en témoignent la traduction que l'on lui attribue d'un texte en vers du Latin Lucrèce et les lettres qu'il a reçues en cette langue, notamment l'éloge versifiée de trois pages que lui avait fait un philosophe de l'époque.

Ensuite, nous avons parlé de la rencontre entre Molière et Corneille à Rouen, selon M. Rey, rien ne prouve que Molière ait rencontré Corneille à cette époque ! En effet, Denis Boissier se base sur la rencontre de Molière avec le notaire Me Cavé en présumant que ce dernier était le notaire de Corneille vu qu'il avait déjà été le notaire de son père mais le nom de Me Cavé ne se lit sur aucun document imprimé concernant Pierre Corneille

En ce qui concerne le manque de temps qu'aurait eu Molière à son arrivé à Paris, François répond qu'il n'en est rien. Aux différents arguments il répond ainsi :

Molière était le directeur d'une troupe des plus importantes de l'époque, avant d'être la plus importante : il n'y avait à l'époque pas de tels « directeurs », l'on peut dire qu'il était le chef, ce qui n'absorbait pas autant de temps que l'on le dit, d'autant plus qu'il partageait la responsabilité de la troupe avec Armande Béjart et Lagrange, acteurs de la troupe.

Molière joue les plus longs rôles et ce tout au long de l'année : certes, mais uniquement pour les comédies ; il est prouvé qu'à cause de ses piètres performances d'acteur tragique, il avait renoncé assez tôt aux rôles dans les tragédies. Il faut remarquer également, grâce aux registres tenus par Lagrange, que les mois précédant la première de *L'Ecole des Femmes*, la troupe ne jouait que des tragédies, ou encore que les douze semaines avant celle d'*Amphytrion*, Molière n'était pas monté sur scène.

Molière était le bouffon du roi : nouvelle trouvaille des « Louÿsiens », ce titre de bouffon était déjà attribué au Marquis de l'Angelis et Molière n'a pas, contrairement à ce que Boissier affirme, les caractéristiques du bouffon : il doit organiser des fêtes, ne fait pas de « bons mots », et d'ailleurs personne à l'époque ne l'a traité ainsi ; il reçoit également des pensions et gratifications attribuées aux gens de lettres, au même titre que Corneille et Racine.

Nous avons ensuite abordé le thème du pseudonyme de Jean-Baptiste Poquelin. Les cornéliens ont émis plusieurs hypothèses auxquelles Monsieur Rey s'est opposé directement : le verbe moliérer serait un synonyme de légitimer, Corneille aurait donc offert ce pseudonyme pour légitimer leur collaboration --> ce verbe est un verbe très ancien, il date des environs du 11ème siècle, et n'est plus usité depuis longtemps à cette époque et définit en fait la légitimation d'un enfant né hors-mariage ; une molière est également une terre détrempée ou une carrière, un pseudonyme féminin pour un homme que les Louÿsiens veulent faire passer pour sodomite. Le moliériste pense plutôt qu'il s'agirait d'un hommage à un certain auteur « de Molière » ayant écrit une œuvre très connue à l'époque et qui a vécu de 1600 à 1624.

Pour finir nous avons parlé de l'étude statistique de Dominique Labbé : il n'a visiblement rien à redire à de telles analyses, mais il est lui frappé par la différence évidente entre la prose des deux hommes (les ressemblances au niveau des textes versifiés en alexandrins sont normales, étant donné la rigueur) : il remarque une surabondance du « et » chez Molière par rapport à Corneille. Il dit même trouver que le style Molière n'est parfois pas très bon, contrairement à celui de son aîné.

Il a enfin conclu en disant que Corneille était bien plus mystérieux que ne l'est Molière, pour lequel de nombreux témoignages de sa vie nous sont parvenus, tels les registres de Grimaret et de Lagrange.

ANNEXE 4 : Exemples et explications du calcul de la distance intertextuelle

Soit trois textes :

A : La chèvre le regarda d'un œil intelligent, leva son pied doré et frappa sept coups. Il était sept heures.

B : Le cabri les regarde d'un air intelligent, lève ses sabots dorés et tape huit coups. Il est huit heures.

C : La gazelle le fixa avec un regard ingénu, leva sa patte frêle et lécha le sol. Elle avait très soif.

Traitement des textes :

On ajoute à chaque mot du texte, une étiquette avec :

- la graphie « standardisée »: « la » pour « La » et « il » pour « Il » (ce qui neutralise la majuscule initiale de phrase),

- une entrée de dictionnaire, ou « vocable », comportant un « lemme » - par exemple « regarder » pour « regarda » ou « il » pour « elle », « son » pour « sa » ou « ses » - et la catégorie grammaticale.

De ce fait « regarda » et « regarde » ou « son », « sa », « ses » sont considérés comme les mêmes vocables (« regarder, v » ou « son, déterminant adjectif possessif »).

Calcul de la distance :

Le calcul est effectué sur les vocables et non pas sur les « mots » (ou « formes graphiques »). Justifications :

- la standardisation est indispensable pour la comparaison des textes entre eux : beaucoup de mots peuvent s'écrire de plusieurs manières, les conventions diffèrent suivant les genres (ex : vers et prose) ;

- la lemmatisation rétablit l'ordre sous-jacent au texte (la langue) tout en correspondant aux habitudes des usagers.

Soit :

- N_a et N_b longueurs de A et B en « mots ». Ici 20 mots dans les trois cas ;

- V_a et V_b le nombre de vocables dans A et B ;

- F_{ia} et F_{ib} le nombre d'apparitions (ou « effectif ») d'un vocable i dans les textes A et B.

$$\text{leur distance relative est égale à : } D_{rel(A,B)} = \frac{\sum_{i \in (A,B)} |V_a - V_b|}{N_a + N_b}$$

Cela donne les tableaux de calcul ci-dessous

Calcul de la distance entre A et B

Vocables	Effectifs A	Effectifs B	Distance (a,b)
le, art	1	1	0
chèvre, n f	1	0	1
cabri, n m	0	1	1
le, pro	1	1	0
regarder, v	1	1	0
de, prep	1	1	0
un, art	1	1	0
œil, nm	1	0	1
air, n m	0	1	1
intelligent, adj	1	1	0
lever, v	1	1	0
son, art	1	1	0
pied, n m	1	0	1
sabot, n m	0	1	1
doré, adj	1	1	0
et, conj	1	1	0
frapper, v	1	0	1
taper, v	0	1	1

sept, num	2	0	2
huit, num	0	2	2
coup n m	1	1	0
il, pro	1	1	0
être, v	1	1	0
heure, n f	1	1	0
Total	20	20	12

La distance relative à :

$$\text{Drel}_{(A, B)} = \frac{\sum_{i \in (A, B)} |V_a - V_b|}{N_a + N_b} = \frac{12}{40} = 0.30$$

Calcul de la distance entre A et C

	A	C	Distance (a,c)
le, art	1	2	1
chèvre, n f	1	0	1
gazelle, n f	0	1	1
le, pro	1	1	0
regarder, v	1	0	1
fixer, v	0	1	1
de, prep	1	0	1
un, art	1	1	0
avec, adv	0	1	1
œil, nm	1	0	1
regard, n m	0	1	1
intelligent, adj	1	0	1
ingénu,adj	0	1	1
lever, v	1	1	0
son, art	1	1	0
pied, n m	1	0	1
patte, n m	0	1	1
doré, adj	1	0	1
frêle, adj	0	1	1
et, conj	1	1	0
frapper, v	1	0	1
lécher, v	0	1	1
sept, num	2	0	2
sol, n m	0	1	1
coup n m	1	0	1
il, pro	1	1	0
être, v	1	0	1
avoir, v	0	1	1
très, adv	0	1	1
heure, n f	1	0	1
soif, n f	0	1	1
Total	20	20	26

Distance relative : 0.65

ANNEXE 5 : Résultats de Labbé sur les textes Gary-Ajar complets

1) Les distances dans le corpus Ajar :

	Gros-Câlin	Vie devant soi	Pseudo	Salomon	Moyenne
Gros câlin	-	0,273	0,224	0,247	0,248
Vie devant soi	0,273	-	0,255	0,176	0,235
Pseudo	0,224	0,255	-	0,246	0,242
SalomonTotal	0,247	0,176	0,246	-	0,223

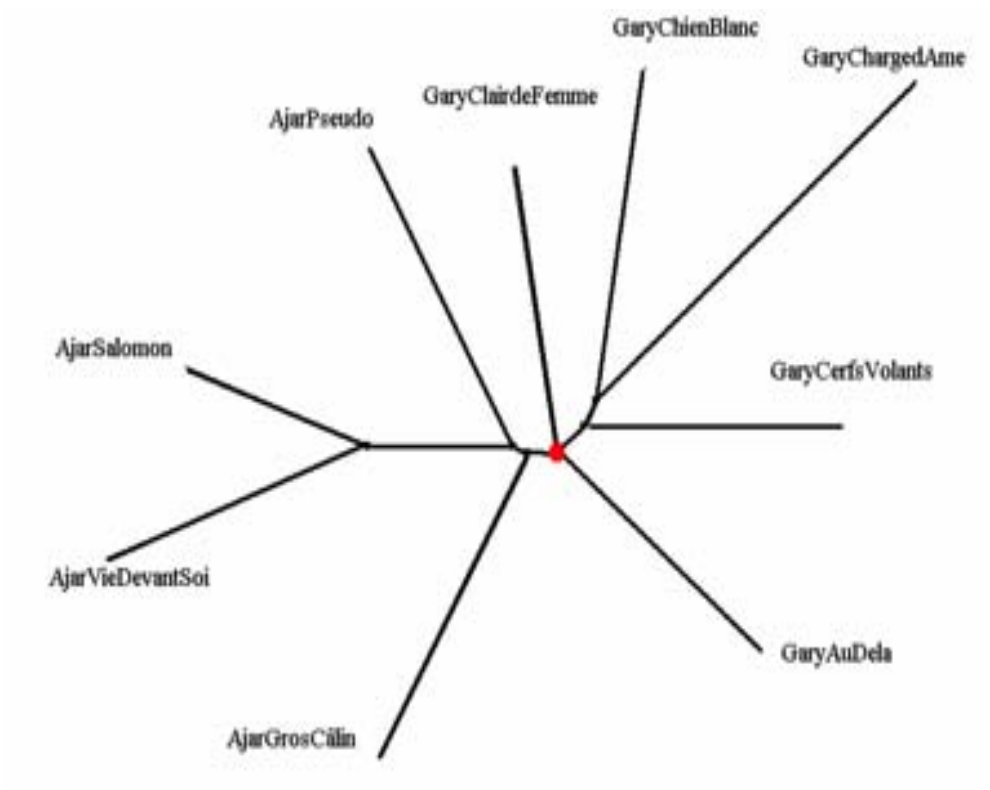
2) Les distances dans le corpus Gary :

	Chien blanc	Au-delà ...	Charge...	Clair de f...	Cerfs-v...	Moyenne
Chien blanc	-	0,227	0,263	0,267	0,224	0,245
Au-delà...	0,227	-	0,280	0,229	0,205	0,235
Charge...	0,263	0,280	-	0,314	0,249	0,276
Clair de f...	0,267	0,229	0,314	-	0,240	0,263
Cerfs-v...	0,224	0,205	0,249	0,249	-	0,230

3) Les distances entre les corpus Ajar et Gary :

	Chien blanc	Au-delà...	Charge...	Clair de f...	Cerfs-v...	Moyenne
Gros Câlin	0,255	0,238	0,311	0,244	0,252	0,260
Vie devant so	0,327	0,310	0,383	0,275	0,309	0,321
Pseudo	0,266	0,247	0,323	0,238	0,256	0,266
SalomonTotal	0,305	0,263	0,360	0,226	0,270	0,284
Moyenne	0,288	0,264	0,344	0,246	0,272	0,283

4) La classification arborée :



BIBLIOGRAPHIE

Introduction: http://insee.fr/fr/ffc/docs_ffc/cs100e.pdf Hamilton/Madison

I/ Sites <http://corneille-moliere.org> Site officiel de l'Affaire Corneille-Molière
<http://www.elitterature.net/publier2/spip/spip.php?article18>

http://fr.wikipedia.org/wiki/Paternit%C3%A9_des_%C5%93uvres_de_Moli%C3%A8re

<http://www.editionsjcgodefroy.fr/page.php?id=90> *L'affaire Molière*, Denis Boissier

Livre *L'Affaire Molière. La grande supercherie littéraire*, de Denis Boissier, 2004

Film *Secrets d'histoire : Molière a-t-il écrit ses pièces?* France 2

II/ Sites <http://corpus.revues.org/index31.html> Explication de la méthode de M. Labbé
<http://members.unine.ch/jacques.savoy/Events/Moliere.html>

Livre *Corneille dans l'ombre de Molière*, de Dominique Labbé, 2003

III/ Sites <http://hal.archives-ouvertes.fr/docs/00/27/96/63/PDF/LabbeGaryAjar2008.pdf>

Livres *La Vie devant soi*, Romain Gary (Emile Ajar) 1975

Clair de femme, Romain Gary, 1977

Conclusion : http://livres.suite101.fr/article.cfm/les_supercheres_litteraires#ixzz0XbyCIEgr
exemples d'affaires

Remarques : Certains de nos écrits se sont appuyés sur des e-mails qui nous étaient adressés.

Le texte en fond de la page de couverture est tiré du *Cid* de Corneille, Acte I scène 1.
Il s'agit d'une tirade d'Elvire.

REMERCIEMENTS

Nous tenons à remercier trois personnes qui nous ont beaucoup aidés, de manière très aimable et qui nous ont ainsi permis de mettre en forme notre projet.

MONSIEUR FRANCOIS REY qui nous a accueillis à son domicile, a répondu à nos questions et nous a éclairés sur différents points au sujet de la vie de Molière et de la thèse cornélienne, de façon très claire et très gentiment.

MONSIEUR DENIS BOISSIER qui a soutenu notre choix de sujet et nous a expliqué différents aspects de la thèse qu'il défend.

MONSIEUR DOMINIQUE LABBE qui nous a consacré beaucoup de son temps afin de nous expliquer ses méthodes de calculs et qui a eu l'amabilité de nous transmettre ses logiciels et d'avoir réalisé pour nous les calculs sur les textes de Gary et Ajar.