



HAL
open science

Lope de Vega, premier polémiste contre Góngora ? Application des méthodes numériques d'attribution d'autorité aux lettres anonymes de 1613-1615

Marie-Eglantine Lescasse

► **To cite this version:**

Marie-Eglantine Lescasse. Lope de Vega, premier polémiste contre Góngora ? Application des méthodes numériques d'attribution d'autorité aux lettres anonymes de 1613-1615. *Observer la vie littéraire. Etudes littéraires et numériques*, 2022. hal-03977731

HAL Id: hal-03977731

<https://hal.science/hal-03977731v1>

Submitted on 7 Feb 2023

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Lope de Vega, premier polémiste contre Góngora ?

Application des méthodes numériques d'attribution d'autorité aux lettres anonymes de 1613-1615

Dans cet article, je me propose d'appliquer différentes méthodes stylométriques d'attribution d'autorité aux lettres anonymes qui, en 1613 ou 1615, ont contribué à lancer la polémique gongorine, en retenant deux candidats : Lope de Vega et Suárez de Figueroa. Les lettres anonymes, au nombre de quatre, proviennent d'un détracteur de Góngora, qui aurait lu les *Soledades*, sans doute dès leur première phase de circulation en 1613. Cette étude se fait dans le cadre du projet Góngora d'édition digitale et d'étude de la polémique, dirigé par Mercedes Blanco, au Labex OBVIL¹.

Nous ne nous proposons pas ici de montrer définitivement l'autorité des lettres anonymes mais, conformément à la méthode stylométrique, d'opposer deux candidats afin d'observer lequel est le plus proche, stylistiquement parlant, du texte à attribuer. Autrement dit, cette étude a moins une valeur positive d'attribution effective qu'une valeur *exploratoire*, en tant qu'expérience numérique permettant de mesurer la probabilité respective que les lettres aient été écrites par chacun des deux auteurs candidats. Il s'agit donc plutôt, en plus de donner des arguments scientifiques pour l'attribution à Lope, de défricher le terrain pour de prochaines études en AA (attribution d'autorité) sur les *cartas anónimas*, en utilisant des méthodes nouvelles de détection du style d'auteur sur des textes espagnols du début du XVII^e siècle, entreprise encore rarement accomplie sur des textes de l'époque classique. Ce n'est donc pas une solution absolue et définitive au problème de l'attribution de ces premières lettres que nous donnerons en conclusion, mais, bien plutôt, les résultats, tout relatifs, de cette étude quantitative et contrastive du style de nos auteurs.

Les lettres anonymes, instigatrices d'une « micro polémique épistolaire »

¹ Voir la page du projet sur le site du Labex OBVIL. – Je tiens également à remercier très vivement Mercedes Blanco pour son aide dans ce travail, commencé en 2017, et pour ses multiples relectures de ce texte.

Contexte

Rappelons d'abord brièvement le contexte d'apparition des lettres : Góngora, soucieux de sonder les opinions pour préparer à son œuvre révolutionnaire le meilleur accueil possible, envoie le *Polifemo* et la *Soledad Primera* à son ami érudit Pedro de Valencia, dans une lettre datée du 11 mai 1613². Il demande à la même époque leur avis à ses amis Francisco Fernández de Córdoba, Abad de Rute, qui rédigera un *Parecer de don Francisco de Córdoba acerca de Las Soledades, a instancia de su autor* (1614), et à l'historiographe Tomás Tamayo de Vargas. Or, cette stratégie de préparation d'une réception favorable se serait accompagnée d'une volonté consciente de rabaisser son rival, Lope de Vega³. Celui-ci, auquel son talent sans égal comme dramaturge et sa fécondité poétique avaient valu le surnom de « Phénix », aurait été froissé par ces échanges entre son ami Tamayo de Vargas et Góngora, ainsi que par la dédicace des *Advertencias* de Almansa y Mendoza au duc de Sessa, ce qui pouvait apparaître comme une provocation, car tout le monde connaissait les liens qu'entretenait cet aristocrate avec Lope, son secrétaire informel et poète attitré.

Ces *Advertencias para inteligencia de las Soledades de don Luis de Gongora* (1613 ou 1614), défense enthousiaste mais peu éclairée de l'œuvre gongorine, sont le fait d'un certain Almansa y Mendoza, personnage de basse extraction, mulâtre et andalou, colporteur de nouvelles à la cour et gazetier, qui serait, selon Henry Ettinghausen et Manuel Borrego qui l'ont édité, le plus grand journaliste du règne de Philippe IV⁴. Il s'agissait de faire face à des attaques qui circulaient déjà à Madrid, où la première *Soledad* passait de main en main sous forme manuscrite, avant même l'envoi à Pedro de Valencia⁵.

En tout cas, il est certain que les *Advertencias* sont antérieures aux lettres anonymes, puisque ces dernières ridiculisent sans cesse l'opuscule et son auteur. La première lettre indique :

*y Andrés de Mendoza se ha señalado en esparcir copias de él. Y no sé si por pretendiente de escribir gracioso, o por otro secreto influjo, se intitula hijo de V.m., haciéndose tan señor de su correspondencia, y de la declaración y publicación de esta poesía*⁶.

Almansa y Mendoza constitue ainsi l'une des principales cibles des lettres, qui s'en prennent à son ignorance⁷. Góngora a-t-il pu confier la défense de son œuvre à un personnage si peu éduqué, et dont il

² Begonia López Bueno, « El cruce epistolar entre Lope y Góngora de 1615-1616. Revisión de fechas », *El Poeta Soledad. Góngora 1609-1615*, Saragosse, Prensas universitarias de Zaragoza, 2011, p. 242.

³ *Ibid.*, p. 248.

⁴ « el más importante periodista conocido de los primeros años del reinado de Felipe IV ». Cités par Begonia López Bueno, « Las advertencias de Almansa y Mendoza, el apócrifo correspondiente de Góngora », *Criticón*, 116, 2012, p. 5-27.

⁵ Begonia López Bueno, « El cruce... », *op. cit.*, p. 243.

⁶ Antonio Carreira, « La controversia en torno a las Soledades. Un parecer desconocido y edición crítica de las primeras cartas », *Hommage à Robert Jammes*, éd. Francis Cerdan, t. 1, Toulouse, Presses universitaires du Mirail, 1994, p. 159. Nous traduisons : « Et je ne sais si pour prétendre au titre d'écrivain burlesque ou à cause de quelque autre influence secrète, il se fait appeler votre fils, se faisant ainsi le maître de votre correspondance et de la publication de votre poésie. » Nous citerons la première lettre et la réponse de Góngora dans l'édition de Carreira, la plus rigoureuse, mais la deuxième et la troisième dans celle d'Emilio Orozco, Carreira ne les ayant pas transcrites dans son article de 1994. Nous avons également consulté Antonio Carreira, *Gongoremas*, Barcelone, Ediciones Península, 1998, p. 239-266.

⁷ « Ésta se ha de verificar con hombres doctos españoles; aquí hay más que en otras partes, que aunque Mendoza los reduce a catorce pudiera acordarse de los Padres Pedrosa, Cerda, de Pedro de Valencia y otros hombres graves y doctos, que [...] los

savait qu'il pourrait faire l'objet d'attaques faciles en raison de ses origines sociales ? Selon Begonia López Bueno, l'acte de Almansa serait indépendant de toute demande de la part du poète⁸ : le gazetier multiplie les erreurs de compréhension de la *Soledad primera* (comme en confondant l'escarboucle ou rubis avec l'animal qui le porte, ou en citant des morceaux de textes viciés) et Góngora l'appelle « *apócrifo correspondiente* » dans une lettre du 4 septembre 1614 à son ami Juan de Villegas.

Quoi qu'il en soit, face aux provocations et au succès croissant de Góngora à la cour, Lope répondit par l'offensive, en rédigeant ou en enjoignant à un ami de rédiger la première lettre, la *Carta escrita a don Luis de Góngora en razón de las Soledades*, à laquelle répond Góngora le 30 septembre 1615, puis une seconde, la *Carta de respuesta*, et une troisième, la *Carta « echadiza »*, donnant naissance à cette « micro polémique épistolaire⁹ », un des épisodes les plus anciens et fondateurs de la polémique dans son ensemble.

Contenu des lettres

Résumons brièvement les lettres. Dans un style incisif et efficace, l'auteur de la première lettre, très courte, a recouru au topos de la confusion linguistique babélique pour ridiculiser le poète des *Soledades* et lui demander de retirer son texte¹⁰.

À cette attaque directe, Góngora répond seulement deux semaines plus tard, par la lettre du 30 septembre 1613 ou 1615, où il expose le principe d'hermétisme élitiste qui gouverne son œuvre¹¹. En déplaçant la signification du mythe de Babel de la production de la parole à sa réception, il procède à un renversement habile de l'argument de la confusion et termine sur une défense du plurilinguisme poétique¹². Robert Jammes a mis en doute que cette lettre soit bien du poète, car elle contient des

que han hecho versos públicos son capaces de materias tan graves », Emilio Orozco Díaz, *op. cit.*, p. 244, où l'auteur anonyme reproche à Mendoza de ne pas citer suffisamment d'autorités à l'appui des *Soledades*. Nous traduisons : « Celle-ci doit s'appuyer sur l'autorité de savants espagnols ; mais en cet endroit il y en a plus que dans d'autres passages, et bien que Almansa y Mendoza les réduise à quatorze, il aurait pu se souvenir des pères Pedrosa, de la Cerda, de Pedro de Valencia, et d'autres hommes sérieux et savants, car ceux qui écrivent des vers pour le public sont capables de matières si sérieuses [...] »

⁸ Begonia López Bueno, « Las Advertencias... », *op. cit.*

⁹ Selon l'appellation de Begonia López Bueno, « El cruce... », *op. cit.*, p. 240-241.

¹⁰ Après un exorde direct et mémorable (« *Un cuaderno de versos desiguales y consonancias erráticas se ha aparecido en esta corte* », « Un cahier de vers inégaux et de rimes erratiques est apparu à la Cour »), dans lequel l'anonyme déguise l'insulte sous le couvert d'une définition exacte, grâce au jeu des syllepthes, entre description poétique et qualifications méprisantes, celui-ci multiplie les louanges feintes (« *apacibilidad del estilo* » - « douceur du style » -, « *pensamientos superiores y donaires agudos* » « pensées élevées et blagues ingénieuses »). L'expression ironique d'un doute sur la langue des *Soledades* (« *Y como ni en éstas ni en las demás lenguas del Calepino están escritos los tales soliloquios* », - « et comme ce n'est ni dans ces langues ni dans les autres langues de Calepino que sont écrits ces soliloques ») se fait dénégation péremptoire des dons linguistiques de l'auteur (« *y se cree que V.m. no ha participado de la gracia de Pentecostés* », - « et il faut croire que vous n'avez pas reçu votre part de l'Esprit de la Pentecôte ») et assimilation railleuse de la langue gongorine à la confusion post-babélique (« *muchos se han persuadido que le alcanzó algún ramalazo de la desdicha de Babel* », - « beaucoup sont persuadés que vous avez été frappé de la punition soudaine de la malédiction de Babel »), et l'auteur enjoint enfin au poète de retirer ses vers. La lettre se trouve dans Antonio Carreira, *art. cit.*, p. 159-161.

¹¹ « *Demás que honra me ha causado hacerme obscuro a los ignorantes, que esa es la distinción de los hombres doctos, hablar de manera que a ellos les parezca griego, pues no se han de dar las piedras preciosas a los animales de cerda* », Antonio Carreira, *art. cit.*, p. 165. Nous traduisons : « En outre, si je me suis rendu obscur aux ignorants, c'est là mon honneur, car c'est là ce qui distingue les doctes : parler de telle sorte que cela leur semble du chinois, car il ne faut pas donner de perles aux cochons. »

¹² « *no los confundió Dios a ellos con darles lenguaje confuso, sino en el mismo suyo ellos se confundieron, tomando tierra por agua y agua por piedra, que esa fue la grandeza de la sabiduría del que confundió aquel soberbio intento. Yo no envió confusas Las Soledades, sino la malicia de las voluntades en su mismo lenguaje halla confusión.* », Antonio Carreira, *art. cit.*, p. 166:

maladresses, tant du point de vue formel que conceptuel, qui ne laissaient d'étonner de la part de don Luis. Un homme d'église pouvait ainsi difficilement attribuer l'épisode de la Pentecôte à l'Ancien Testament, et les paragraphes IV à VIII, les plus douteux et écrits en un style plus lourd, auraient pu être interpolés par un admirateur¹³. Antonio Carreira, de son côté, défendit l'attribution de la lettre dans son entier à Góngora, tout en précisant, non sans humour, qu'il n'aurait « jamais dû l'écrire¹⁴ ».

La seconde lettre anonyme façonne une manière de fiction épistolaire. Son auteur se présente comme un intermédiaire chargé de transmettre une réponse au nom d'un ami soldat parti à Naples, qui serait l'auteur de la première lettre à laquelle répliquait Góngora. Après mille précautions et demandes d'indulgence, l'auteur accuse le poète d'avoir attaqué son ami brutalement, de détruire la langue avec son nouveau style, et lui conseille de se remettre sérieusement en cause.

L'auteur de la troisième et dernière lettre anonyme, appelée par Emilio Orozco « *Carta echadiza* », se présente comme un lettré portugais diplômé de l'université de Coimbra. Cette posture distante lui permet de se livrer à un déferlement d'injures adressées à Mendoza, dans les litanies burlesques de l'exorde, qui dépeignent le facteur de Góngora en pantin de cour agité, cherchant la compagnie des puissants, puis dans la comparaison avec le Pasquin de Rome. Après avoir dépeint Almansa y Mendoza, « *Paraninfo de los predicadores* », « *Mercurio de las nuevas y sátiras de este Reino*¹⁵ », comme le pire défenseur possible (pire que Jáuregui aurait pu l'être !), l'auteur, qui affirme bien connaître Lope (« *clérigo y vecino mío* », « cleric et mon voisin »), rejette la faute sur l'intermédiaire en affirmant qu'il se serait trompé : Lope n'aurait jamais fait montre d'animadversion contre Góngora, et toute cette affaire ne serait qu'un terrible malentendu, Góngora cherchant à déshonorer un poète qui ne fait que s'appliquer qu'à le louer¹⁶.

Comme le rappelle Roland Béhar dans son article sur les polémiques littéraires de l'époque moderne¹⁷, ces procédés de fictionnalisation, typiques des querelles poétiques classiques, permettaient de transformer le conflit, régi par un ensemble de codes et des règles à respecter, à la manière d'un duel, en œuvre littéraire, mise en scène se conformant elle-même aux règles de l'écriture satirique antique, reprises au XVI^e siècle en Italie par les grands polémistes que furent Pietro Aretino et Francesco Berni.

« Dieu ne les a pas confondus en leur donnant une langue confuse, mais ils se sont confondus eux-mêmes dans leur propre langue, nommant l'eau terre et eau la pierre, et là fut la grandeur du miracle. Pour ma part, mes *Soledades* ne sont pas confuses, mais c'est la malveillance des lecteurs qui leur fait trouver de la confusion dans leur propre langue » et plus bas : « no van en más que en una lengua *Las Soledades*, aunque pudiera, quedándome el brazo sano, hacer una miscelánea de griego, latino y toscano con mi lengua natural, y creo no fuera condenable; que el mundo está satisfecho que los años de estudio que he gastado en varias lenguas han aprovechado algo a mi corto talento », « les *Soledades* ne sont pas écrites en plus d'une langue, même si j'aurais pu, et j'aurais encore eu de la marge, faire un mélange de grec, de latin et de toscan avec ma langue naturelle, ce qui à mon avis n'aurait pas été condamnable, pour que le monde soit convaincu que mes années d'études en langues ont profité à mon pauvre génie ».

¹³ Antonio Carreira, *art. cit.*, p. 168-169.

¹⁴ Antonio Carreira, *art. cit.*, p. 171.

¹⁵ Emilio Orozco Díaz, *op. cit.*, p. 263.

¹⁶ « De suerte que todo su estudio de V. m. es solicitar el deshonor de este hombre, y todo el suyo celebrar su ingenio de V. m. », Emilio Orozco Díaz, *op. cit.*, p. 266.

¹⁷ Roland Béhar, « Homeromastix, Vergiliomastix... ¿Gongoramastix? », *e-Spania*, 18 juin 2014, paragraphe 1, en ligne, <https://journals.openedition.org/e-spania/23769>.

Du point de vue stylistique, la première lettre se distingue fortement des deux autres, par son style incisif, sa brièveté, son humour direct et efficace et sa plus grande simplicité syntaxique (moins de subordinées). Les lettres II et III présentent des phrases longues, faites de multiples détours rhétoriques et d'une pléiade de propositions subordinées.

Témoignages et datation

Les témoignages existants de la première lettre ont été étudiés par Antonio Carreira¹⁸, et leur multiplicité ainsi que leur datation incertaine a contribué à rendre l'attribution difficile. Sans entrer à fond dans ces questions, il semble néanmoins nécessaire de rappeler l'état des textes. Il existe cinq témoignages manuscrits de la première lettre anonyme et de la réponse de Góngora, regroupés par Carreira en deux familles, les deux assez contaminées : la famille α , constituée du manuscrit 3811 de la BN, transcrit par Paz y Melia en 1902, le manuscrit 612 de la bibliothèque publique municipale de Porto, et le Barberini Latini 3476, de la Bibliothèque vaticane, découvert par le dominicain Maximiliano Canal Gómez en 1931. Ce dernier codex aurait appartenu au cardinal Barberini et contient, en plus des lettres situées à la fin, des poèmes de Góngora. La famille β comprend un manuscrit ayant appartenu au duc de Gor, étudié par Emilio Orozco Díaz dans *En torno a las Soledades* (1969), aujourd'hui conservé à la Bibliothèque de la Fondation Bartolomé March à Palma de Majorque. Il existe aussi, dans cette même bibliothèque, et provenant du même fond, un codex intitulé *Escritos curiosos. Cartas de varones doctos de España. Papeles originales. Letras y firmas de diferentes eruditos*, ayant appartenu à Juan Iriarte puis à sir Thomas Phillips. Ce manuscrit contient la lettre I, la réponse de Góngora, et la lettre II.

Selon Robert Jammes et Antonio Carreira, les lettres dateraient des débuts de la polémique. L'un des arguments importants en faveur d'une écriture précoce, en plus de l'impression que donne l'exorde d'une réaction à chaud (« *Un cuaderno de [...] consonancias erráticas se ha aparecido en esta corte* »), serait qu'Almansa utilise une version primitive des *Soledades* – ce qu'ont montré Roses et Jammes¹⁹ –, antérieure à celle qu'aurait reçu Pedro de Valencia, proche du manuscrit rédigé à Cordoue pendant l'écriture de la première *Soledad*, retrouvé par Rodríguez Moñino. Selon Begoña López Bueno, la présence de plusieurs poèmes en défense des *Soledades* au milieu des textes du recueil, datés de 1615 selon le manuscrit Chacón, signifierait une rédaction plus tardive, en 1615, et il faudrait alors revoir notre conception de la chronologie des débuts de la polémique²⁰.

¹⁸ Antonio Carreira, *art. cit.*

¹⁹ Begoña López Bueno, « Las advertencias... », *op. cit.*

²⁰ Begoña López Bueno, « El cruce epistolar... », *op. cit.*, p. 240, note 1.

Justification du choix des auteurs-candidats – Premiers arguments pour l’attribution à Lope

Quoi qu’il en soit de leur datation, ces lettres ont été attribuées à Lope, depuis leur première impression par Antonio Paz y Meliá en 1902. Par la suite, cette attribution a été reprise par Raymond Fouché-Delbosc, Juan et Isabel Millé, Joaquín de Entrambasaguas, puis Emilio Orozco Díaz dans *Lope y Góngora frente a frente* (1973), sur lequel nous nous appuyons ici, et n’a jamais vraiment été remise en cause. La rivalité entre Lope de Vega et Góngora, qui s’envenime avec l’arrivée à la cour de ce dernier en 1617, existait dès le début des années 1610, comme le montre Emilio Orozco. Du point de vue philologique, les arguments sont à mon avis nombreux pour retenir l’autorité de Lope, en particulier pour la lettre II. Nous nous bornerons ici à les lister afin de passer rapidement à l’analyse automatique.

La mention de Lipse dans la lettre II rappelle la référence récurrente au grand érudit dans l’œuvre de Lope, qui eut l’idée un peu curieuse de considérer Góngora comme un Lipse de la poésie. D’après lui, si le nouveau latin de Lipse a quelque chose d’admirable, celui de ses imitateurs est exécration. Il en va de même pour le castillan de Góngora. C’est ce qu’il fait dans la censure de la nouvelle poésie qu’il imprime sous son nom dans *La Filomena*, et, dans les deux fragments, une métaphore tératologique lui est appliquée²¹.

Les deux textes contiennent également le fantasme menaçant d’une destruction de la langue, dans la mise en scène d’un complot des « *cultos* » contre le castillan dans *La Filomena*, et l’emploi de *destruir* dans la lettre II²².

²¹ Comparer, dans *La Filomena* : « *Lipsio escribió aquel nuevo latín de que dicen los que le saben que se han reído Cicerón y Quintiliano en el otro mundo, y siendo tan doctos los que le han imitado, se han perdido; y yo conozco alguno que ha inventado otra lengua y estilo tan diferente del que Lipsio enseña, que podía hacer un diccionario, como los ciegos a la jergonza: y así, los que imitan a este caballero producen partos monstruosos que salen de generación, pues piensan que han de llegar a su ingenio por imitar su estilo* », Félix Lope de Vega, *Epístolas de La Filomena de Lope de Vega*, édition critique par Pedro Conde Parrado, 2015, en ligne sur le site de l’OBVIL, <https://urlz.fr/eDXe>, f° 194r^o. Nous traduisons : « Lipse a écrit ce latin étonnant dont on dit que Cicéron et Quintilien se sont moqués outre-tombe, et bien que ses imitateurs fussent très doctes, ils se sont ainsi discrédités ; quant à moi, je connais quelqu’un qui a inventé une autre langue et un style si différent de celui de Lipse, qu’il pourrait en faire un dictionnaire, comme les aveugles en font de leur jargon, de telle sorte que les imitateurs de ce gentilhomme accouchent d’œuvres monstrueuses et dégénérées, pensant pouvoir atteindre son génie par l’imitation de son style. » Et dans la lettre II : « *siempre alaba y encarece aquel género de transposiciones en su elegante poesía de V.m., y consta a toda la gente que le hizo por algunos mochuelos que aquí le imitan bárbara y atrevidamente, a quien sucede lo que a muchos que contrahacen el latín de Justo Lipsio y escriben una lengua tan monstruosa que ni es latina, ni hebrea, ni arábica, mas no por esto Lipsio deja de ser aquel divino inventor de tan único estilo, que es lo mismo que sucede a V.m., único ingenio y inimitable* », Emilio Orozco Díaz, *op. cit.*, p. 264-265. Nous traduisons : « il [Lope] loue et exalte toujours ce genre d’hyperbates dans votre élégante poésie, et il est évident pour tout le monde qu’il l’a fait [écrire un sonnet infamant sur Góngora] à cause de quelques rapaces qui vous imitent en cela de manière téméraire et barbare, et à qui arrive ce qui arrive à beaucoup d’autres imitateurs du latin de Juste Lipse, qui écrivent dans une langue si monstrueuse qu’elle n’est ni latine, ni hébraïque, ni arabe ; cependant Lipse ne cesse pas d’être pour autant le divin inventeur de ce style si unique, et il en va de même pour vous et votre génie, unique et inimitable. » La mention du latin de Juste Lipse, invention nouvelle, le jargon incompréhensible (« *jergonza* » / « *que ni es latina, ni hebrea, ni arábica* ») qui naît de son imitation (« *los que imitan* » / « *algunos mochuelos que aquí le imitan* »), et la monstruosité de la langue hybride qui en résulte (« *partos monstruosos* » / « *lengua tan monstruosa* ») apparaissent dans les deux passages. Même la structure des fragments est proche, dans la chute sur le génie inimitable des deux auteurs (« *piensan que han de llegar a su ingenio por imitar su estilo* » / « *único ingenio e inimitable* »).

²² « *Esta excelencia, siendo de las mayores de nuestra lengua, la destruye V.m. con su nueva gramática: mire qué lejos está de perfeccionarse con ella.* », Emilio Orozco Díaz, *op. cit.*, p. 244. Nous traduisons : « Cette excellence, qui pourtant est l’une des plus importantes de notre langue, vous la détruisez avec votre nouvelle grammaire : voyez combien vous êtes loin de vous perfectionner en castillan ».

Le déploiement d'érudition de la lettre II, où l'auteur évoque des anecdotes antiques de Socrate, Alquidamas, Philippe et l'ambassadeur Agis, rappelle celui dont fait preuve Lope dans *La Filomena*, où les citations érudites interviennent à des moments de prise de risque polémique²³. Cet étalement de connaissances, associé au maniement de phrases très longues, à la syntaxe compliquée, laissent imaginer que Lope se plaît à « noyer le poisson » en faisant se succéder digressions savantes à visée lénifiante et piques soudaines, qui n'en sont que plus violentes.

On sait aussi que Lope avait une profonde aversion pour Almansa y Mendoza, qui lui inspirait dégoût et agacement : il le compare à un chien, suprême insulte que l'on réservait à l'époque aux morisques, dans sa lettre du 8 mai 1615, où il dit l'avoir rencontré à Zocodover²⁴.

Enfin, dans la lettre n° 295 de l'édition de González de Amezúa adressée au duc de Sessa, légèrement postérieure à la précédente, Lope mentionne une lettre qu'il aurait écrite à Góngora ou à son sujet²⁵.

Pour ce qui est de Suárez de Figueroa, sa position, extrêmement ambiguë, n'est jamais explicitée dans son œuvre.

Roland Béhar, dans son article sur les polémiques poétiques espagnoles modernes, développe plusieurs arguments en faveur du positionnement proGóngora de Suárez de Figueroa : il aurait été l'ami de Pedro de Valencia, érudit et défenseur de l'œuvre gongorine, ainsi que de Manuel Ponce²⁶. Mais, si l'on étudie de plus près le cercle de relations de Suárez de Figueroa, il n'en irait pas forcément ainsi : l'auteur était un ami intime de Pedro de Torres Rámila, maître de grammaire à l'université d'Alcalá à partir de 1613 et principal auteur de la virulente attaque contre Lope en 1617, la *Spongia*, libelle en latin destiné à effacer par absorption – d'où la métaphore spongiaire –, les œuvres déjà parues du Phénix, qui ne respectaient pas les principes sacro-saints de la poétique aristotélicienne. Les deux amis, ainsi qu'un troisième larron, Juan Pablo Mártir Rizo, également polémiste contre Lope, se réunissaient, encore étudiants, dans la *tertulia* de la librairie du français Jerónimo de Courbés à Madrid, située calle Mayor, en face des marches du couvent de San Felipe el Real²⁷. D'après Julián González-Barrera, ce serait même sur les conseils de Suárez de Figueroa que Torres Rámila aurait entrepris son œuvre d'éreintage

²³ Par exemple, après la suggestion sous-entendue de retirer son œuvre, Lope cite tour à tour Aulu Gelle, Cyprien, Quintilien, Pic de la Mirandole, Héraclite, Platon, et Mercure Trismégiste. Félix Lope de Vega, *Epístolas, op. cit.*, f° 193r° à 194r°.

²⁴ « *En Zocodover me asió la mano Mendoza; pensé que me la quería morder, y cubrila con el manto; no quiero yo decir en esto que es perro, sino que lo es de las bodas, pues se halla hasta en las de los quemados* », Félix Lope de Vega, *Epistolario de Lope de Vega Carpio que por acuerdo de la Real Academia Española publica Agustín G. de Amezúa*, éd. Agustín González de Amezúa y Mayo, Madrid, Artes Gráficas Aldus S.A., 1941-1943, p.187. « À Zocodover, Almansa y Mendoza me saisit la main, je crus qu'il voulait me la mordre et la couvrir de ma cape ; je ne veux pas dire par là que c'est un chien, mais qu'il l'est des mariages, puisqu'il se rend même aux noces des brûlés », allusion à un autodafé auquel assista Almansa y Mendoza.

²⁵ « *Las cartas verá Vuestra excelencia, Señor: la de Salinas, a propósito, y las otras, al tiempo; la que trata de Góngora no se ha copiado, que es de tres pliegos y no tengo oficial más de mi pluma, cuando no ocupada, falta de buena letra* », Félix Lope de Vega, *op. cit.*, p. 286. Nous traduisons : « Vous verrez les lettres : celle à Salinas, bien à propos et les autres, quand vous voudrez ; celle qui porte sur Góngora n'a pas été recopiée, car elle fait dix pages et je n'ai pas de greffier autre que ma propre plume, dépourvue d'une belle écriture, lorsqu'elle n'est pas déjà occupée. »

²⁶ Roland Béhar, *art. cit.*, paragraphe 12.

²⁷ Xavier Tubau Moreu, *Lope de Vega y las polémicas literarias de su época. Pedro de Torres Rámila y Diego de Colmenares*, thèse doctorale, Universidad Autónoma de Barcelona, 2008, en ligne, <https://urlz.fr/eDXo>, p. 27, n. 10.

de Lope²⁸. Suárez de Figueroa aurait lui-même participé à l'écriture de la *Spongia*, et s'en serait pris, en une autre occasion, au mécène et ami de Lope, le duc de Sessa²⁹.

Or, la situation au sein du champ de la polémique est plus compliquée. Si Suárez de Figueroa se plaçait sans doute effectivement du côté des « néo-aristotéliens³⁰ », avec Juan de Jáuregui dont il est l'ami, et Cristóbal de Mesa, contre Lope, cette position n'entraîne pas nécessairement, par transitivité, une inscription dans le camp gongorin. Ainsi, d'une part, si le *Certamen poético* d'octobre 1616 pour l'inauguration de la chapelle de Nuestra Señora del Sagrario à Tolède amène le triomphe de Góngora et, d'après Álvarez Amo, une alliance entre les « cultistes » et les « néo-aristotéliens » contre Lope, cela n'empêche pas ces derniers de s'en prendre aux imitateurs de Góngora³¹ – mais reste à déterminer si leur critique s'applique également à Góngora lui-même. D'autre part, et comme l'écrit très justement Álvarez Amo, les polémiques littéraires étaient les lieux de tous les retournements, et la fidélité à l'un ne signifiait pas nécessairement la prise de parti pour son ennemi. Non seulement les positions sont poreuses et changent au gré des événements et des modifications du champ de la polémique, mais encore, la complexité des positions individuelles ne se laisse pas résumer à l'appartenance à tel ou tel camp, et il convient ainsi de tenir compte des affinités interpersonnelles, aptes à transcender les querelles de groupe³². L'hypothèse d'un Suárez de Figueroa opposé à Góngora ne semble donc pas à écarter sur la seule base de son inimitié manifeste pour Lope de Vega.

Regardons plutôt ce qui justifie, dans l'œuvre du polémiste, de figurer parmi nos deux candidats : intéressons-nous autrement dit aux indices d'une critique de l'auteur à l'encontre de Góngora. Nous verrons que les arguments de Suárez de Figueroa contre notre poète sont très semblables à ceux de l'auteur anonyme³³.

Tout d'abord, si l'auteur loue le génie poétique du Cordouan, qu'il appelle « *luciente honor de las Españas* » dans le *Pusílipo* (1629), et s'il lui reconnaît le talent d'un Martial dans ses vers de *burla* (vers

²⁸ Julián González-Barrera, « Expostulatio Spongiae redux. Aristóteles versus Horacio », *AnMal*, XXXIV, 2, 2011, p. 497-516, p. 498.

²⁹ Francisco Javier Álvarez Amo, « «Yo no presumo de poeta...». La trayectoria de Juan de Jáuregui », *Etiópicas*, 7, 2011, p. 120-138, p. 135.

³⁰ Nous reprenons l'expression de Francisco Javier Álvarez Amo, *art. cit.*, p. 133.

³¹ Francisco Javier Álvarez Amo, *art. cit.*, p. 134.

³² À propos de la trajectoire de Jáuregui: « *La trayectoria de Jáuregui ilustra, más bien, el carácter sistemático de la república de las Letras en el Siglo de Oro. El valor de las posiciones individuales depende de, y repercute sobre, el valor de la posición del resto de individuos. Cualquier alteración del campo da lugar a casi imprevisibles efectos de acción y reacción. [...] La trayectoria de Jáuregui deja, de esta forma, de contener misteriosos e inexplicables virajes a favor y en contra de Luis de Góngora o Lope de Vega: sus acciones y omisiones, en efecto, como las del resto de sus contemporáneos, como las de los nuestros, nacen del conflicto entre sus aspiraciones individuales y las coordinadas poéticas de su tiempo.* », Francisco Javier Álvarez Amo, *art. cit.*, p. 138. Nous traduisons: « La trajectoire de Jáuregui illustre plutôt le caractère systématique de la république des Lettres au Siècle d'or. La valeur des positions individuelles dépend de, et se répercute sur la valeur de la position du reste des individus. Toute altération du champ donne lieu à des effets d'action et de réaction quasiment imprévisibles. [...] La trajectoire de Jáuregui cesse, de cette manière, de présenter de mystérieux et inexplicables retournements en faveur ou en défaveur de Góngora ou de Lope : ses actions et ses omissions, en effet, comme celles du reste de ses contemporains, comme celles des nôtres, naissent du conflit entre ses aspirations individuelles et les coordonnées poétiques de son temps. »

³³ Nous tenons à remercier José María Rico pour son aide au début de cette étude, et de nous avoir mis sur la piste d'une possible attribution à Suárez de Figueroa.

burlesques), c'est plutôt en Stace qu'il le dépeint du côté des *veras* (vers sérieux)³⁴. Ce dernier, poète latin du I^{er} siècle et auteur des *Silves*, était, à l'époque de Góngora, l'un des auteurs antiques les plus admirés, cités et imités, l'une des grandes autorités de la Rome impériale, que seuls surpassaient Virgile et Horace. Réputé pour sa grande difficulté, c'est donc par hyperbole que les polémistes le comparent à Góngora, afin d'exagérer l'obscurité de notre poète, comme le fait également Jáuregui dans son *Discurso poético* (1621)³⁵.

Par ailleurs, le polémiste discret s'en prend aux principes esthétiques de Góngora dans un passage bien connu du *Pasajero* (1617). Dans un fragment de l'« *Alivio II* », l'auteur fustige les poètes qui introduisent des mots archaïques, latins et des néologismes dans leur œuvre. Il les compare alors, comme dans la deuxième lettre anonyme, à Merlin Cocaye, pseudonyme burlesque de l'humaniste italien Teofilo Folengo, auteur du *Baldus* (1517-1521) et instigateur dans sa Macaronée – surnom donné à son œuvre – du latin macaronique, hybride de latin et de dialectes toscans :

Sin duda, se levanta en España nueva torre de Babel, pues comienza a reinar tanto la confusión entre los arquitectos y peones de la pluma. No sirve el hablar de encubrir o poner en tinieblas los conceptos, sino de descubrirlos y declararlos. Merlin Cocayo, donosísimo poeta, aludiendo en su Macarronea a este lenguaje infernal, introduce a un demonio hablando, sin poder ser entendido, de esta manera: Drum Cararontardus, tragaron granbeira detronde. El Dante, por el consiguiente, varón doctísimo, hace en su obra que Lucifer, admirado de ver en su región hombres en carne y hueso, exclame diabólicamente: Pape Satán, Pape Satán, Alepe³⁶.

Après la référence à Babel, qui rappelle l'attaque de la première lettre anonyme, l'auteur se moque de l'obscurité poétique des nouveaux Cocayes, caricature déjà utilisée plusieurs fois par Quevedo. La citation, supposément tirée de Folengo, en langue inventée, fonctionne comme satire de la « langue mêlée » gongorine, supposément faite d'un galimatias d'autres langues, en aucun cas castillane. La citation en langue démoniaque contribue à opacifier le langage, l'éloignant ainsi de sa fonction première de communication transparente, pour la réduire à une suite de signes vains, métaphorisant une langue gongorine où l'hypertrophie des signifiants bloque tout accès au signifié, en attirant toute l'attention sur la forme, devenue première et autosuffisante, au détriment du fond.

³⁴ Dans son éloge de Góngora inclus dans le *Pusilipo* : « *Este es aquel monstruo de los ingenios, aquel fénix de las agudezas, don Luis de Góngora: el Solo poeta español, el moderno Marcial, más que él agudo en las burlas, y, en las veras, otro Papinio Estacio* », Cristóbal Suárez de Figueroa, *Pusilipo. Ratos de conversación en los que dura el paseo*, édition numérique de Jonathan Bradbury, en ligne sur le site de l'OBVIL, <https://urlz.fr/eDXr>, p. 260. Nous traduisons : « C'est lui de tous les génies le plus prodigieux, le phénix de la pointe, don Luis de Góngora, le seul et unique Poète espagnol, le Martial moderne, plus spirituel que celui-ci dans les poèmes comiques tandis que, dans les poèmes sérieux, il égale Stace. » Encore une fois, cet éloge tardif ne montre en rien dans quel camp pouvait se trouver Suárez de Figueroa au début de la polémique.

³⁵ Voir Cristóbal Suárez de Figueroa, *op. cit.*, n. 6.

³⁶ Cristóbal Suárez de Figueroa, *El Pasajero*, édition électronique de Jonathan Bradbury, 2016, en ligne sur le site de l'OBVIL, <https://urlz.fr/eDXs>. Nous traduisons : « Sans doute est-on en train de construire en Espagne une nouvelle Tour de Babel, tant commence à régner la confusion parmi les architectes et ouvriers de la plume. La parole ne sert pourtant pas à voiler ou à entourer les idées de ténèbres, mais à les révéler et à les énoncer clairement. Merlin Cocaye, poète très amusant, faisant allusion dans la Macaronée à ce langage infernal, introduit un démon parlant de cette manière, et que l'on ne peut comprendre : *Drum Cararontardus, tragaron granbeira detronde*. Dante, également, homme très docte, montre dans son œuvre Lucifer, étonné de voir dans sa région des hommes en chair et en os, s'exclamer de manière diabolique : *Pape Satán, Pape Satán, Alepe*. »

Or, d'après Jonathan Bradbury dans ses notes à l'édition numérique du *Pasajero*, si Suárez de Figueroa pouvait également penser dans l'« *Alivio II* » au théoricien cordouan Luis Carrillo y Sotomayor, défenseur dans le *Libro de la erudición poética* (1611) d'une poétique élitiste et « hermétiste », que notre polémiste a paradoxalement imité à de nombreuses reprises dans sa propre poésie³⁷, la référence au « premier auteur » de la nouvelle poésie ne pouvait s'appliquer qu'à Góngora : c'est donc bien, incontestablement, Góngora lui-même qui constitue la cible de la critique dans ce passage³⁸.

L'année de publication du *Pasajero*, 1617, est d'ailleurs celle de l'arrivée de Góngora à la Cour madrilène et de la publication de la *Spongia* contre Lope, culmination de la rivalité entre les deux poètes et *annus horribilis* pour le Phénix³⁹. La formulation claire et explicite d'une critique antigongorine à une année décisive de la polémique ne laisse pas de nous incliner à penser à une prise de position de Suárez de Figueroa contre Góngora⁴⁰.

Face à ces multiples arguments d'ordre philologique, les méthodes numériques d'attribution d'autorité peuvent-elles nous aider à résoudre ce cas épineux ?

³⁷ Voir Ángeles Arce, *Cristóbal Suárez de Figueroa. Nuevas perspectivas de su actividad literaria*, thèse doctorale, Madrid, Editorial de la Universidad, 1983, édition électronique de 2015, en ligne, <https://urlz.fr/eDXw>, p. 70-90.

³⁸ Cristóbal Suárez de Figueroa, *op. cit.*, n. 3.

³⁹ Selon la jolie expression de Francisco Javier Álvarez Amo, *art. cit.*, p. 134.

⁴⁰ Enfin, mais cet argument ne nous semble pas suffisant à lui seul pour prouver la réticence de Suárez de Figueroa à l'encontre de Góngora, comme l'a remarqué María José Osuna Cabezas, c'est très vraisemblablement à la satire d'Almansa y Mendoza que se livre Suárez de Figueroa dans l'« *Alivio IV* » du *Pasajero*, en ces termes : « *No es tan veloz el rayo como sus pies para dar con ligerísima ocasión una vuelta al mundo. [...] Si le tuviérades por amigo, pudiérades a ojos cerrados ocupar el púlpito, y aun estoy por decir osar predicar sin meditación, casi de repente. Subiera vuestro nombre a las nubes, exagerara pomposamente vuestras letras, y esparciera vuestras alabanzas con tan resonantes hipóboles y encarecimientos, que no hicieran tanta operación si todas las hojas de los árboles fueran lenguas, si todas las arenas del mar fueran voces. Ignora totalmente los primeros rudimentos latinos, mas encomienda a la memoria con tan grande puntualidad las autoridades de Escritura y Evangelios, que deja asombrados la primera vez que le oyen a los más entendidos, juzgándole por extremo erudito en letras humanas. Su prosa es redundante y hueca. [...] Válese de exquisitas palabras: condensar, retroceder, equiparar, asunto y otras así. Huye cuanto puede los términos humildes, siguiendo cierta afectación ostentativa.* » Nous traduisons : « Même la foudre n'est pas aussi vélocité que ses pieds pour faire le tour du monde. [...] Si vous l'aviez comme ami, vous pourriez occuper le pupitre les yeux fermés, et même, j'allais dire, oser prêcher sans préparation, presque au pied levé. Il porterait votre nom aux nues, il exagérerait vos lettres avec pompe, et il répandrait vos louanges avec des hyperboles et des insinuations si grandiloquentes, que toutes les feuilles des arbres n'en feraient pas autant si elles étaient des langues, ni tous les grains de sable de la mer s'ils étaient des voix. Il ignore totalement les premiers rudiments du latin, mais il rappelle à la mémoire les autorités de la Bible et des Évangiles avec tant d'exactitude qu'il laisse surpris les plus doctes, la première fois qu'ils l'entendent, qui le jugent extrêmement érudit en lettres humaines. Sa prose est redondante et creuse. [...] Il se sert de mots choisis : *condensar, retroceder, equiparar, asunto*, et d'autres du même acabit. Il évite autant qu'il le peut les termes simples, suivant en cela une certaine affectation ostentatoire. » Le passage ridiculise Almansa y Mendoza, et serait le signe de l'inimitié de Suárez de Figueroa pour le colporteur. Mais, encore une fois, un dégoût individuel ne laisse en rien présager d'un positionnement au sein d'un champ collectif. Voir María José Osuna Cabezas, « El papel de Andrés de Almansa y Mendoza en la polémica gongorina », *Edad de oro cantabrigense. Actas del VII Congreso de la Asociación Internacional Siglo de Oro*, AISO, Robinson College, Cambridge, 18-22 juillet 2005, éd. Anthony Close, p. 489-494, p. 490-491.

Analyse quantitative avec les méthodes de l'attribution d'autorité (AA)

AA et stylométrie

L'attribution d'autorité constitue, avec la caractérisation stylistique, une des deux applications principales de la stylométrie. Celle-ci désigne l'étude du style auctorial par des mesures statistiques sur des traits jugés distinctifs et appelés « marqueurs stylistiques ». Il s'agit donc d'un ensemble de techniques de quantification du style, surtout employées pour la reconnaissance d'auteur⁴¹.

D'une manière générale, les traits utilisés sont soit de nature lexicale (fréquence et richesse du vocabulaire, observation des hapax) soit de nature syntaxique (fréquence des POS-tags ou des n-grams de POS-tags). Jack Grieve, et plus récemment Efstathios Stamatatos, listent respectivement l'ensemble des traits sur lesquels peut porter l'analyse : longueur des mots, des phrases, richesse du vocabulaire, fréquence de graphèmes, de signes de ponctuation ; *most frequent words*, n-grams de mots, n-grams de POS-tags, n-grams de caractères⁴².

Jusqu'aux années 1990, une préférence pour les indicateurs de type lexical était répandue en AA⁴³. Holmes écrivait encore en 1994 : « *To date, no stylometrist has managed to establish a methodology which is better able to capture the style of a text than that based on lexical items*⁴⁴ », en entendant par *lexical items* des mots de contenu.

On préfère désormais s'intéresser aux mots-outils, c'est-à-dire les pronoms, articles, prépositions, qui présentent l'avantage d'être très fréquents dans les textes.

Les mots-outils, une révolution dans la conception du « style » d'auteur

Ce glissement de l'objet d'étude de traits peu fréquents à des marqueurs très fréquents a révolutionné la manière dont on envisage le style, et est couramment associé au travail fondamental de Mosteller et Wallace sur les *Federalist Papers* – une série d'essais politiques écrits par John Jay, Alexander Hamilton et James Madison – (1964), comme le note Mike Kestemont :

⁴¹ Efstathios Stamatatos, « A survey of modern authorship attribution methods », *Journal of the American society for information science and technology*, num. 60, 3, en ligne, <https://urlz.fr/eDWL>.

⁴² Jack Grieve, « Quantitative Authorship Attribution. An evaluation of techniques », *Literary and Linguistic Computing*, vol. 22, num. 3, 2007, en ligne, <https://urlz.fr/eDWN>. Et E. Stamatatos, *op. cit.*

⁴³ C'est ce que Stamatatos appelle les *traditional authorship attribution studies*, par opposition au tournant des *non traditional AA studies*. Efstathios Stamatatos, *op. cit.*, p. 538.

⁴⁴ Cité par Mike Kestemont, « Function words in authorship attribution. From black magic to theory? », *Proceedings of the 3rd Workshop on Computational Linguistics for Literature*, 2014, p. 62. Nous traduisons : « Jusqu'à ce jour, aucun chercheur en stylométrie n'a réussi à trouver une méthode qui saisisse mieux le style d'un texte que les méthodes basées sur le lexique. »

*The fundamental novelty of the work by Mosteller and Wallace was that they advised to move away from a language's low-frequency features to a language's high-frequency features, which often tend to be function words*⁴⁵.

Paradoxalement, le style ne se situerait pas au niveau de singularités, d'hapax, de bizarreries, *whimsicalities*⁴⁶ qui semblent pourtant hautement personnelles et identifient un individu auteur pour le lecteur, mais au niveau de ce qu'il y a de plus commun, de plus partagé, de plus « évident », d'apparemment moins personnel : les mots-outils. Si l'appellation mots-outils ou *function words* est très discutable en ce qu'elle est couramment associée à une idée de vacuité sémantique qui pourrait être remise en question (il y a une sémantique de l'information grammaticale)⁴⁷, leur utilisation en AA s'est révélée extrêmement performante ces dernières années et ceux-ci sont à présent considérés comme un des marqueurs stylistiques les plus discriminants⁴⁸. Les linguistes sont eux-mêmes divisés sur la question du sémantisme des mots-outils, même si la plupart s'accorde à les qualifier d'asémantiques. Ainsi, Nathalie Rossi-Gensane et Laurice Tuller écrivaient-elles, citant à leur tour André Martinet :

Selon André Martinet, « il existe un rapport constant et inverse entre la fréquence d'une unité et l'information qu'elle apporte » ou, autrement dit, « plus une unité est fréquente, moins elle est informative ». On ne peut alors accorder qu'un sens limité aux unités de cette sorte, que l'on considère que ces unités n'ont pas de sens – ou en ont trop peu – ou que l'on considère, au contraire, qu'elles ont trop de sens, c'est-à-dire de trop nombreux sens, le recours au contexte étant nécessaire pour sélectionner l'un d'entre eux⁴⁹.

Les mots-outils auxquels les linguistes trouvent plus volontiers une « coloration » sémantique sont les prépositions, où l'on peut distinguer, en suivant Spang-Hanssen, les prépositions « incolores » ou vides de sens, des prépositions « colorées » ou sémantiquement fortes⁵⁰. Mais ce n'est pas ici notre propos.

Pour revenir au choix des mots-outils comme marqueurs stylistiques, Mike Kestemont a relevé la similarité du déplacement des *content words* aux *function words* avec celle de motifs iconographiques reconnus à des détails apparemment insignifiants en histoire de l'art⁵¹ : Giovanni Morelli attribuait des peintures à des maîtres sur la seule base de la forme des mains, des oreilles, des pieds⁵². Par ailleurs,

⁴⁵ *Ibid.*, p. 60. Nous traduisons : « La nouveauté fondamentale du travail de Mosteller et Wallace fut de conseiller de passer de l'étude des traits linguistiques de faible fréquence à l'étude des traits très fréquents, ce qui revient la plupart du temps à étudier les mots-outils. »

⁴⁶ Selon le mot de Mike Kestemont, « Function words... », *op. cit.*, p. 60.

⁴⁷ Voir Philippe Bourdin, « Venir en français contemporain. De deux fonctionnements périphrastiques », *Les périphrases verbales*, éd. Hava Bat-Zeev Shyldkrot et Nicole Le Querler, Philadelphie, J. Benjamins, 2005, p. 261.

⁴⁸ Mike Kestemont, « Function words... » *op. cit.*

⁴⁹ Nathalie Rossi-Gensane et Laurice Tuller, « Introduction », *Les grammaires du français et les mots-outils*, Actes du colloque international de linguistique tenu à Tours les 29 et 30 juin 1998, Gudrun Ledegen et Nathalie Rossi-Gensane (dir.), Caen, Presses Universitaires de Caen, 2002, p. 17-18.

⁵⁰ Nathalie Rossi-Gensane, « Pour une valeur aspectuelle des prépositions. Le cas des prépositions en emploi spatial dynamique », *Ibid.*, p. 132-133.

⁵¹ *Ibid.*, p. 61.

⁵² On pourrait comparer ces motifs apparemment anodins, marqueurs d'autorité, aux signatures stylistiques que constituent les suites d'étiquettes grammaticales. Nous avons essayé dans un autre travail de détecter ces « traces » grammaticales de l'écriture gongorine, définissant à la fois une appartenance auctoriale et des stylèmes récurrents et interprétables : Marie-Églantine

l'emploi des mots-outils serait largement inconscient, ou peu contrôlé de la part de l'auteur. Depuis plusieurs années, cette idée circulait dans la communauté stylométrique, mais elle était rarement justifiée⁵³: Argamon et Levitan relient seulement leur emploi irréfléchi à leur forte fréquence et à leur statut de mots grammaticalisés⁵⁴, Stamatatos se contente d'affirmer que leur caractère incontrôlable rend les *function words* d'excellents marqueurs du style d'auteur⁵⁵.

Le style se placerait ainsi du côté de la nature plutôt que de l'art et ne serait pas vraiment imitable. Involontairement et à son insu, l'auteur se trahirait dans le style ainsi conçu, d'où l'idée de travailler sur ces marqueurs afin de démasquer un auteur qui se dissimule ou cherche à se faire passer pour un autre, comme c'est le cas dans la polémique gongorine. Cette approche permet également de se départir des préjugés critiques qui font que l'on aimerait attribuer un texte à tel ou tel. L'AA s'assimile à une enquête policière destinée à piéger un coupable. Dans un article lumineux traduit en français en 1986, le grand historien Carlo Ginzburg s'intéressait ainsi à ce qu'il appelle un « paradigme indiciaire », et qu'il applique à l'analyse picturale⁵⁶. Il y décrivait la méthode de Giovanni Morelli dans les années 1870 sur le modèle du détective qui découvre l'auteur d'un délit par des indices imperceptibles : certains détails caractéristiques trahiraient un auteur comme un criminel est trahi par ses empreintes digitales. Mais surtout, il reliait ces « faits marginaux révélateurs » à la part inconsciente de l'art, aux moments où le contrôle de l'artiste se relâche pour laisser place à des traits purement individuels. Il s'agit de « chercher la personnalité là où l'effort personnel est le moins intense⁵⁷ » et de réaliser ainsi une « identification du noyau intime de l'individualité artistique aux éléments soustraits au contrôle de la conscience⁵⁸ ».

Or, cette conception du style, laissant une très large place au rôle de l'inconscient, est à rapprocher de l'analyse que fait Roland Barthes du style dans « Qu'est-ce que l'écriture⁵⁹ ? » Même si le passage est très long, nous aimerions le citer ici *in extenso* :

Le style est presque au-delà [de la Littérature] : des images, un débit, un lexique naissent du corps et du passé de l'écrivain et deviennent peu à peu les automatismes mêmes de son art. Ainsi sous le nom de style, se forme un langage autarcique qui ne plonge que dans la mythologie personnelle et secrète de l'auteur, dans cette hypophysique de la parole, où se forme le premier couple des mots et des choses, où s'installent une fois pour toutes les grands thèmes verbaux de son existence. Quel que soit son raffinement, le style a toujours quelque chose de brut : il est une forme sans destination, il est le produit d'une poussée, non d'une intention, il est comme une dimension verticale et solitaire de la pensée. Ses références sont au niveau d'une biologie ou d'un passé, non d'une Histoire : il est la « chose » de l'écrivain, sa splendeur et sa prison, il est sa solitude. Indifférent et transparent à la société, démarche close de la personne, il n'est nullement le produit d'un choix, d'une réflexion sur la Littérature. Il est la

Lescasse, « Góngora hors norme ? Étude stylométrique d'un motif gongorin », *Theorien von Autorschaft und Stil in Bewegung. Romanische Studien, Beihefte*, éd. Nanette Rißler-Pipka et Kai Nonnenmacher, Regensburg, 2019, p. 91-116.

⁵³ Mike Kestemont, « Function words... », *op. cit.*, p. 61.

⁵⁴ Shlomo Argamon et Shlomo Levitan, « Measuring the usefulness of function words for authorship attribution », *Proceedings of ACH/ALLC Conference*, Janvier 2005, p. 1. En ligne, <https://urlz.fr/eDWQ>.

⁵⁵ Efstathios Stamatatos, *op. cit.*, p. 5.

⁵⁶ Carlo Ginzburg, « Traces. Racines d'un paradigme indiciaire », *Mythes, emblèmes, traces. Morphologie et histoire*, Lagrasse, Verdier, 2010, p. 139-180.

⁵⁷ Carlo Ginzburg, *art. cit.*, p. 225.

⁵⁸ Carlo Ginzburg, *art. cit.*, respectivement p. 225 et p. 231.

⁵⁹ Roland Barthes, « Qu'est-ce que l'écriture ? », *Le degré zéro de l'écriture*, Paris, Seuil, 1972, p. 15-20.

part privée du rituel, il s'élève à partir des profondeurs mythiques de l'écrivain, et s'éploie hors de sa responsabilité. Il est la voix décorative d'une chair inconnue et secrète ; il fonctionne à la façon d'une Nécessité, comme si, dans cette espèce de poussée florale, le style n'était que le terme d'une métamorphose aveugle et obstinée, partie d'un infra-langage qui s'élabore à la limite de la chair et du monde. Le style est proprement un phénomène d'ordre germinatif, il est la transmutation d'une Humeur⁶⁰.

Dans ces pages à la fois lumineuses et cryptiques, sibyllines et solaires, le critique se livre à une véritable biologisation du style, qui en fait moins l'expression d'une pensée et d'un travail sur la langue, qu'un affleurement du corps et du moi intime dans l'écriture, une « poussée » et une germination sur laquelle l'auteur n'aurait aucune prise. Déterminisme physiologique, le style est le destin de l'écrivain (une « Nécessité »), quelque chose d'inévitable, dont il ne pourra jamais se libérer. Si Barthes semble plutôt penser aux marqueurs lexicaux dans cette analyse (« un lexique na[ît] du corps et du passé de l'écrivain » ; « le premier couple des mots et des choses »), celle-ci s'applique cependant parfaitement aux découvertes actuelles de la stylométrie sur l'importance de l'inconscient stylistique, que le grand critique anticipait de soixante-dix ans.

Mike Kestemont a ainsi montré que le rôle clef des mots-outils en AA se justifie à la lumière de la psycholinguistique ou linguistique cognitive, dont plusieurs études⁶¹ ont prouvé que l'emploi – en tout cas la lecture ou la réception – de mots-outils était inconsciente, ceux-ci ne faisant pas l'objet d'une attention particulière de la part du récepteur, qui les considère comme évidents (« *taken for granted* ») ; ils reçoivent un traitement cognitif automatique. Or ils sont aussi omniprésents et structurent le discours. Le choix des MFW (*most frequent words*) – donc des mots outils – comme traits de l'analyse, participant de ce renversement dans la conception du style, du conscient à l'inconscient, comporte en outre un certain nombre d'avantages, listés par José Nilo Binongo puis par Mike Kestemont⁶²:

- Ils sont incontournables, les auteurs d'une même époque sont obligés d'utiliser dans une même langue les mêmes mots-outils, ce qui accroît la comparabilité des textes.
- Leur haute fréquence est un avantage pour les calculs statistiques, dont la significativité dépend de la taille de l'échantillon.
- Ils sont censés ne pas dépendre du genre ni du thème d'une œuvre. Cet aspect est en fait critiqué depuis quelques années, car certains types de mots fonctionnels, comme les pronoms, sont très liés au genre (par exemple le pronom de 1^{re} personne au genre autobiographique, de 3^e personne au roman, etc)⁶³.
- Comme nous venons de le voir, leur emploi est plus inconscient, moins contrôlé, que celui des autres mots, ils permettent donc de prendre un auteur « en flagrant délit » d'autorité.

⁶⁰ *Ibid.*, p. 16.

⁶¹ Mike Kestemont cite Schindler, Drewnowski & Healy, Aronoff & Fudeman, McCusker, Gough & Bias, Morrow : Mike Kestemont, *op. cit.*, p. 61 *sq.* On pourrait ajouter Friederici sur l'automatisme du traitement des mots-outils : Angela Friederici, « Children's sensitivity to function words during sentence comprehension », *Linguistics*, 21, 1983, p. 717-739.

⁶² José Nilo Binongo, « Who wrote the 15th book of Oz? an application of multivariate analysis to authorship attribution », *Chance*, vol. 16, num. 2, p. 9-17. Et Mike Kestemont, *op. cit.*

⁶³ Mike Kestemont, *op. cit.*, p. 61.

On travaillera donc sur les 100 MFW, paramètre par défaut de Stylo R, mais aussi celui qui montre la meilleure performance.

L'AA, définition, méthodes et métriques

Quant à l'attribution d'autorité à proprement parler, il s'agit de l'usage des méthodes de classification textuelle afin d'identifier l'auteur d'un texte anonyme ou d'autorité douteuse. Le chercheur grec Efstathios Stamatatos la définit justement comme une méthode de catégorisation :

In the typical authorship attribution problem, a text of unknown authorship is assigned to one candidate author, given a set of candidate authors for whom text samples of undisputed authorship are available. From a machine learning point-of-view, this can be viewed as a multi-class single-label text categorization task (Sebastiani, 2002). This task is also called authorship (or author) identification usually by researchers with a background in computer science⁶⁴.

Il s'agit donc de chercher à classer un texte, dans un groupe ou *cluster* de textes similaires, selon des critères définis, à partir d'un ensemble fini d'auteurs candidats. On appelle ce processus, visant à réunir dans des groupes homogènes des données (ou individus) hétérogènes, le *clustering* ou partitionnement de données.

Les méthodes et les métriques pour la classification sont nombreuses : du côté des méthodes, elles se divisent en approche supervisée et non supervisée. Parmi les méthodes non supervisées, on peut utiliser le *clustering*, l'analyse en composantes principales (ACP), ou le positionnement multidimensionnel (MDS). Les méthodes de l'intelligence artificielle avec corpus d'apprentissage les plus utilisées en AA sont les k plus proches voisins (KNN), les machines à vecteurs de support (SVM), la classification naïve bayésienne (*Naive Bayes*) et les plus proches centroïdes (NSC), qui sont tous des algorithmes de classification. La métrique, c'est-à-dire le calcul de distance utilisé, varie elle aussi, selon que l'on choisisse le « Delta classique » ou Delta de Burrows⁶⁵, celui d'Eder, la distance de Manhattan (somme des valeurs absolues des différences entre coordonnées), de Canberra (Manhattan pondéré). Les mesures de distance s'appliquent aux traits choisis pour l'analyse : si l'on travaille sur les 100 MFW (*Most Frequent Words*), l'algorithme calcule la distance (c'est-à-dire soustrait) des fréquences des formes⁶⁶ de ces 100 mots chez un auteur par rapport à l'autre. On calcule ainsi des vecteurs de formes,

⁶⁴ Efstathios Stamatatos, *op. cit.*, p. 2. Nous traduisons : « Dans un problème typique d'attribution d'autorité, on assigne un texte d'autorité inconnue à un auteur-candidat, à partir d'un ensemble de candidats pour lesquels on dispose d'échantillons de texte d'autorité incontestée. Du point de vue de l'apprentissage-machine, on peut voir cela comme une tâche de classification en classes multiples (Sebastiani, 2002). Cette tâche est également appelée identification d'auteur par des chercheurs venant du champ des sciences de l'information. »

⁶⁵ L'appellation « delta » correspond à « différence », étant donné qu'il s'agit de soustractions entre fréquences de formes ou d'étiquettes.

⁶⁶ Rappelons que la « forme » en statistique textuelle désigne l'archétype d'occurrences identiques, tandis que les occurrences désignent leurs différentes actualisations dans le texte. Voir André Salem et Ludovic Lebart, *Statistique textuelle*, Paris, Dunod, 1994, p. 313 et p. 315.

dont l'ensemble contribue à situer un texte sur un graphique. La distance entre deux auteurs se visualise dès lors sous forme spatiale, bien que la distance calculée soit de nature non-euclidienne.

Outils et corpus utilisés dans cette étude

Logiciels

Les outils que nous allons utiliser pour cette étude sont Stylo R, un package de R développé à l'université de Cracovie par Maciej Eder, Mike Kestemont et Jan Rybicki⁶⁷, et TXM, logiciel de textométrie développé à Lyon par Bénédicte Pincemin, Serge Heiden, Alexei Lavrentiev, Matthieu Decorde⁶⁸. Stylo R a connu un vaste succès ses dernières années auprès de la communauté des chercheurs en stylométrie, en raison de sa performance et de l'ergonomie de son interface graphique, très facile d'utilisation. Pour paramétrer comme on le souhaite les analyses, il faut cependant utiliser des lignes de code, assez intuitives elles aussi.

Malgré ce succès, il a été assez peu utilisé sur des corpus du Siècle d'or espagnol. Les pionniers en ce domaine ont été Nanette Rißler-Pipka, qui a étudié l'autorité du *Quichotte* apocryphe⁶⁹ et Laura Hernández Lorenzo, dans sa thèse sur les poèmes d'attribution incertaine de Fernando de Herrera⁷⁰. Nous nous proposons donc d'utiliser cet outil pour tenter de découvrir l'autorité des lettres anonymes qui ont lancé la polémique gongorine. Réussir à attribuer ces textes constituerait une véritable révolution dans notre manière d'envisager cet épisode : parce qu'elles appartiennent à la toute première phase de la polémique, ce sont des pièces essentielles de ce qui sera l'une des plus grandes querelles littéraires espagnoles, et, peut-être, européennes du XVII^e siècle⁷¹.

Pour ce qui est des méthodes, de la métrique, des traits observés et des paramètres, l'ensemble des possibilités offertes par Stylo sera exploré, à la fois via l'interface graphique et en lignes de commande, pour obtenir une vue d'ensemble, probabiliste, sur notre corpus. Toutes ces possibilités seront détaillées au cours de l'article.

⁶⁷ Maciej Eder, Jan Rybicki et Mike Kestemont, « Stylometry with R: a package for computational text analysis », *The R Journal*, vol. 8, 1, août 2016.

⁶⁸ Voir le site du projet Textométrie.

⁶⁹ Nanette Rißler-Pipka, « Avellaneda y los problemas de la identificación del autor. Propuestas para una investigación con nuevas herramientas digitales. », *El Otro Don Quijote. La continuación de Fernández de Avellaneda y sus efectos*, éd. Hanno Ehrlicher, Eichstätt, Katholische Universität Eichstätt, Zentralinstitut für Lateinamerika Studien, Mesa redonda nueva serie, 2016, num. 33, p. 27-54.

⁷⁰ Thèse intitulée *Los textos poéticos de Fernando de Herrera. Aproximaciones desde la estilística de corpus y la estilometría*, soutenue en janvier 2020 et pas encore publiée.

⁷¹ Voir l'introduction de Mercedes Blanco à la page OBVIL du projet Góngora, « Góngora et les querelles littéraires de la Renaissance », en ligne, <https://urlz.fr/eDXC>.

Le corpus

Pour réaliser cette étude, j'ai passé à l'OCR⁷² la totalité de l'œuvre en prose de Lope de Vega, en utilisant l'édition de Donald McGrady de 1997-1998⁷³, soit :

- Les romans :
 - *La Arcadia* (1598) ;
 - *El peregrino en su patria* (1604) ;
 - *Los pastores de Belén* (1612).
- Les nouvelles à Marcia Leonarda (1624) :
- *Novela de La Filomena : Las fortunas de Diana* ;
- *Novelas de la Circe* :
 - *La desdicha por la honra* ;
 - *La prudente venganza* ;
 - *Guzmán el Bueno*.
- *La Dorotea* (1632).
- Les lettres au duc de Sessa de 1609⁷⁴.

Je n'ai retenu qu'une partie des lettres de Lope pour ne pas introduire de biais dans l'analyse, du fait de la disparité numérique entre l'énorme masse de la correspondance du Phénix (à peu près 35 900 mots, entièrement ocrisés), et les très rares lettres qui nous sont parvenues de Suárez de Figueroa (4278 mots). Pour ce qui est de Suárez de Figueroa, nous disposons de :

- *El Pasajero. Advertencias utilísimas a la vida humana* (1617).
- *La Plaza Universal de todas las ciencias y artes* (1615).
- *Pusilipo. Ratos de conversación en los que dura el paseo* (1629)⁷⁵.
- La correspondance, issue de l'édition de J. P. Wickersham Crawford, in : *Vida y obras de Cristóbal Suárez de Figueroa*⁷⁶:
 - *Memorial al gobernador de Milán*, 15 sept. 1597 ;
 - *Cartas al duque de Mantua relativas a El Pastor Fido*, 21 déc. 1608, 14 mars 1609, et 22 sept. 1609 ;
 - *Carta con motivo de su destitución del cargo de auditor en Lecce*, 22 août 1624.

⁷² J'ai utilisé le logiciel ABBY Fine Reader Professional.

⁷³ Félix Lope de Vega, *Prosa I et Prosa II*, édition critique par Donald McGrady, Madrid, Fundación José Antonio de Castro, 1997-1998.

⁷⁴ À partir de l'édition suivante de la correspondance : Félix Lope de Vega, *Epistolario de Lope de Vega Carpio que por acuerdo de la Real Academia Española publica Agustín G. de Amezúa*, éd. Agustín González de Amezúa y Mayo, Madrid, Artes Gráficas Aldus S.A., 1941-1943.

⁷⁵ Tous trois pris des éditions en ligne de Enrique Suárez Figaredo de 2004-2005, <https://urlz.fr/eDWZ>, <https://urlz.fr/eDX7>, <https://urlz.fr/eDXb>.

⁷⁶ Maniée dans la réédition en ligne de Enrique Suárez Figaredo de 2005, <https://urlz.fr/eDX4>.

Cela nous donne en tout : 324 596 mots du côté de Lope de Vega, 583 514 du côté de Suárez de Figueroa. Les trois lettres anonymes ont été océrisées à partir de l'édition d'Amezúa de la correspondance de Lope⁷⁷ :

- *Carta escrita a don Luis de Góngora en razón de las Soledades* (650 mots) ;
- *Carta de respuesta* (4896 mots) ;
- *Carta « echadiza »* (1478 mots)⁷⁸.
-

Nous avons donc en tout 8110 mots pour les trois lettres. Le corpus est donc très hétérogène, à la fois par une variabilité de la taille des textes intra-auteur (par exemple la *Dorotea* fait 33 102 mots et la nouvelle *La desdicha por la honra* 5405) mais aussi interauteurs (le corpus de Suárez de Figueroa fait presque le double de celui de Lope), entre les textes à attribuer et les textes d'autorités connues, entre les textes anonymes eux-mêmes (la seconde lettre fait presque quatre fois la longueur de la dernière, et huit fois celle de la première). Pour se rendre compte de l'hétérogénéité du corpus, voici le tableau du nombre de mots par texte :

Texte	Taille
Figú_cartas	2526
Figú_pasajero	14178
Figú_plaza	148837
Figú_Pusílipo	40552
Lettre_ano_1	394
Lettre_ano_2	2965
Lettre_ano_3	931
Lope_arcadia	29845
Lope_cartas	2750
Lope_desdicha	5405
Lope_diana	6045
Lope_dorotea	33102
Lope_guzman	5588
Lope_pastores	31422
Lope_peregrino	29633

Tableau 1 – Nombre de mots par texte

⁷⁷ Félix Lope de Vega, *Epistolario... op. cit.*

⁷⁸ Pour toutes les œuvres, nous retiendrons dans la suite de l'étude un titre raccourci conventionnel (par exemple *Desdicha* pour *La desdicha por la honra*).

Deux individus se distinguent fortement, avec une longueur de 50 000 à 100 000 mots pour le premier, atteignant 150 000 mots pour le second. Il s'agit du *Pasajero* et de la *Plaza universal* de Suárez de Figueroa. Faut-il alors les enlever ? Cela introduirait un fort déséquilibre dans l'ensemble, qui ne contiendrait plus qu'un seul texte en prose non épistolaire de Suárez de Figueroa, le *Pusílipo*. On choisit donc de conserver le *Pasajero* et la *Plaza*.

Problèmes liés au corpus

À la vue de ces données, on identifie deux problèmes importants qu'il faudra prendre en compte dans le choix des méthodes utilisées : la disparité de taille entre les éléments du corpus, qui devrait être palliée par le travail sur des fréquences relatives, et, plus grave, la taille très réduite de certains textes, dont deux des trois lettres anonymes. La taille minimale des textes à attribuer a fait l'objet de débats au sein de la communauté des stylométriciens : si Holmes proposait en 2001 la limite de 1000 mots⁷⁹, la tendance est depuis à la hausse, et Burrows proposait un an plus tard la taille de 1500⁸⁰. Dans un article récent qui mérite de faire consensus, Eder établit une fourchette de 2500 à 5000 mots, avec une performance accrue sur des extraits plus longs⁸¹.

L'hétérogénéité générique pose aussi problème, mais le choix d'étudier les mots les plus fréquents devrait réduire l'importance du genre dans le *clustering*.

Prétraitements

Avant de lancer nos analyses, on a pondéré les fréquences lexicales par la taille des textes, afin de travailler sur des fréquences relatives.

Sur ce corpus, on a réalisé de multiples tests, avec des méthodes non supervisées dans un premier temps, supervisées dans un second temps. Ces dernières sont des méthodes de classification fondées sur l'intelligence artificielle avec l'utilisation de corpus d'apprentissages.

⁷⁹ David Holmes et Richard Forsyth, « Feature-finding for text classification », *Literary and Linguistic computing*, 11, 4, 1996, p. 163-174, p. 167.

⁸⁰ John Burrows, « 'Delta': A measure of stylistic difference and a guide to likely authorship », *Literary and Linguistic Computing*, num. 17, vol. 3, 2002, p. 267-287.

⁸¹ Maciej Eder, « Does size matter? Authorship attribution, small samples, big problem », *Digital Humanities 2010 Conference Abstracts*, King's College London, 2010, p. 132-135, p. 132.

Méthode utilisée : approches non supervisées, et premiers résultats

Résultats de l'ACP

La première approche a été l'ACP ou analyse en composantes principales (PCA en anglais)⁸². Il s'agit d'une méthode d'analyse statistique multidimensionnelle consistant à réduire la dimensionnalité d'un ensemble de données, par exemple dans le cas de tableaux de données à entrées multiples appelés tableaux croisés ou tables de contingence. Dans ce cas, on possède un grand nombre de variables quantitatives fonctionnant comme descripteurs des individus de la population statistique observée. L'idée est qu'une trop forte dimensionnalité dans les données rend leur interprétation difficile ; l'analyse consiste à la fois à réduire la quantité d'informations en l'homogénéisant, et à représenter graphiquement les résultats. Il s'agit ensuite de projeter les résultats sur une série de plans factoriels, ou plans orthogonaux en deux dimensions maximisant l'inertie entre les points du nuage, c'est-à-dire leur dispersion. Le but est ainsi de discriminer les individus entre eux, d'observer des regroupements et des distinctions, que l'on appelle « attraction » et « répulsion ». Concrètement, en lexicométrie et stylométrie, cela permet de repérer des individus (textes) présentant un profil lexical similaire, c'est-à-dire présentant les mêmes mots dans les mêmes proportions. Mais on peut bien sûr lancer l'analyse sur de nombreux autres facteurs observables : cooccurrences de mots, n-grams de caractères, de POS-tags, etc. Les similarités et différences seront alors mesurées d'après le facteur que l'on aura retenu comme critère discriminant. Les deux graphiques suivants ont été générés par l'analyse de Stylo R sur les 100 MFW des textes de notre corpus décrit plus haut, les textes anonymes ne recevant aucun traitement particulier.

⁸² Nous nous appuyons pour les paragraphes suivants sur l'ouvrage de Brigitte Escofier et Jérôme Pagès, *Analyses factorielles simples et multiples. Cours et études de cas*, Paris, Dunod, 2016, p. 7-34.

Individuals factor map (PCA)

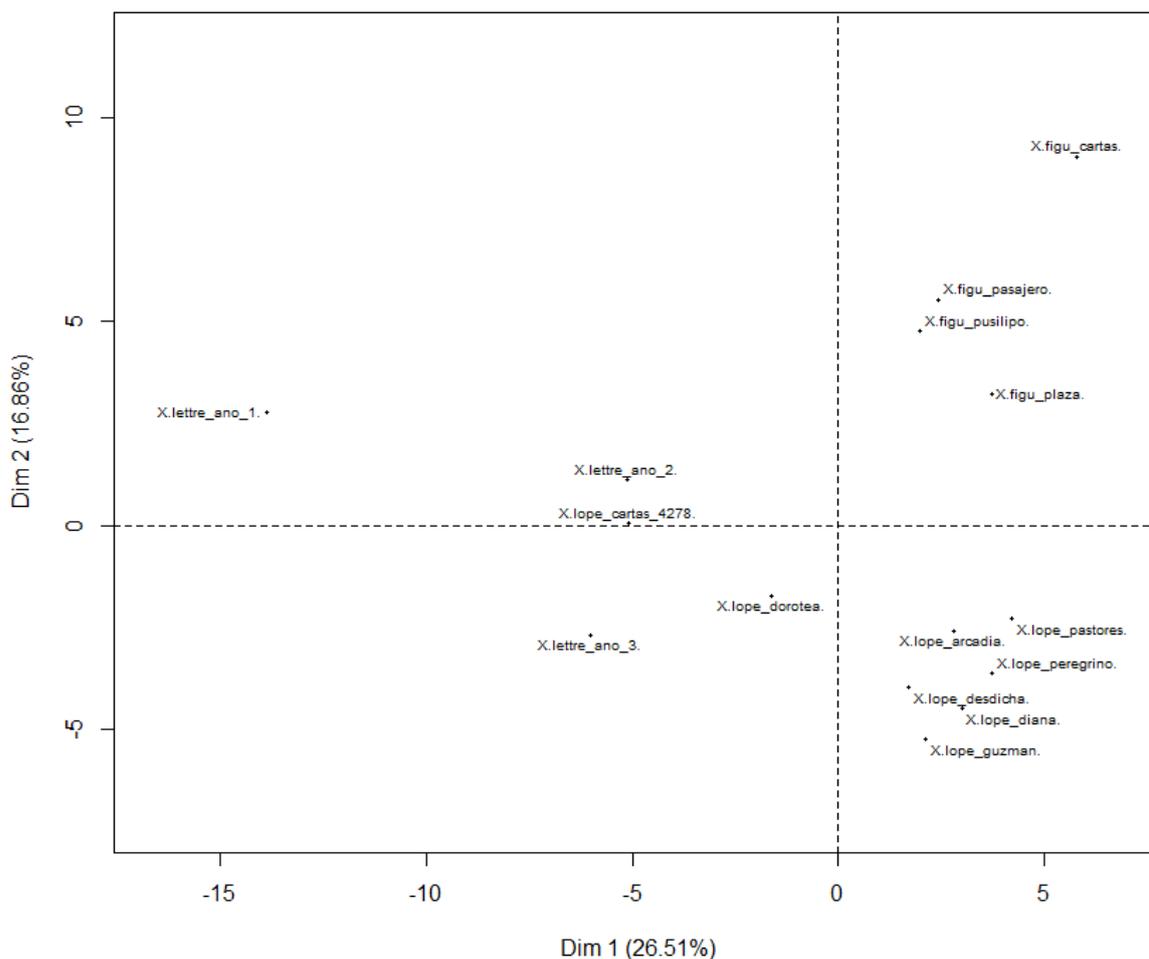


Figure 1 – Axes factoriels de l'ACP

Les deux axes les plus significatifs sont les axes 1 et 2, représentés graphiquement par les lignes en pointillés, avec des valeurs propres de respectivement 26,5 et 16,8. Les « valeurs propres » ou mesures d'inertie des axes ou facteurs correspondent à une mesure de dispersion des points du nuage projetés sur les axes. En ACP, on cherche à maximiser l'inertie afin de mieux discriminer les textes entre eux, autrement dit, plus la valeur propre d'un axe est élevée, plus cet axe disperse les points du nuage (les textes) et donc mieux il permet de les distinguer entre eux⁸³.

Les axes factoriels se lisent d'abord indépendamment l'un de l'autre, selon les dimensions horizontales puis verticales, en commençant par l'axe le plus discriminant, puis par cadran, l'origine du repère, ou barycentre, représentant le profil moyen. L'éloignement d'un point du nuage au barycentre signifie que le texte en question possède un vocabulaire qui lui est propre, et fortement distinct des autres.

⁸³ Voir Brigitte Escofier et Jérôme Pagès, *op. cit.*, p. 23 pour un résumé de la méthodologie suivie en ACP, p. 386 et p. 389 pour les définitions des termes « inertie » et « valeur propre ».

Ici, la séparation entre droite et gauche de l'axe 1 semble pouvoir se lire selon l'opposition générique entre épistolaire et romanesque. L'apparition de la *Dorotea* à gauche de l'axe évoque une opposition en termes de niveaux de style : les lettres et ce roman dialogué, très oral, se distinguent par un usage des pronoms personnels plus important que dans les romans, avec beaucoup de marques de première et de deuxième personne. Il est étonnant que les lettres de Suárez de Figueroa se situent à droite de l'axe, mais cela est peut-être dû à la brièveté du corpus épistolaire de l'auteur. La PCA semble reconnaître l'écriture de cet auteur, tous genres confondus, en l'isolant dans le cadran haut-droit ; elle distingue ensuite le Lope romancier (cadran droite bas : *Pastores, Arcadia, Peregrino*) du Lope épistolier (cadrans de gauche). On peut déjà noter la forte proximité entre la lettre II et la correspondance de Lope.

La méthode de l'ACP consiste à réduire la dimensionnalité en la ramenant à une série de dimensions non corrélées entre elles. Que se passerait-il si, plutôt que d'observer les deux premiers axes de dix dimensions – c'est-à-dire les deux axes aux valeurs propres les plus élevées, les deux axes les plus discriminants –, on demandait d'emblée à l'algorithme de réduire les données à deux dimensions ? On code donc l'obtention d'un graphique en deux dimensions, résumant la totalité de l'information de la table lexicale de départ, grâce à la méthode du positionnement multidimensionnel ou *Multidimensional Scaling*.

Résultats du positionnement multidimensionnel

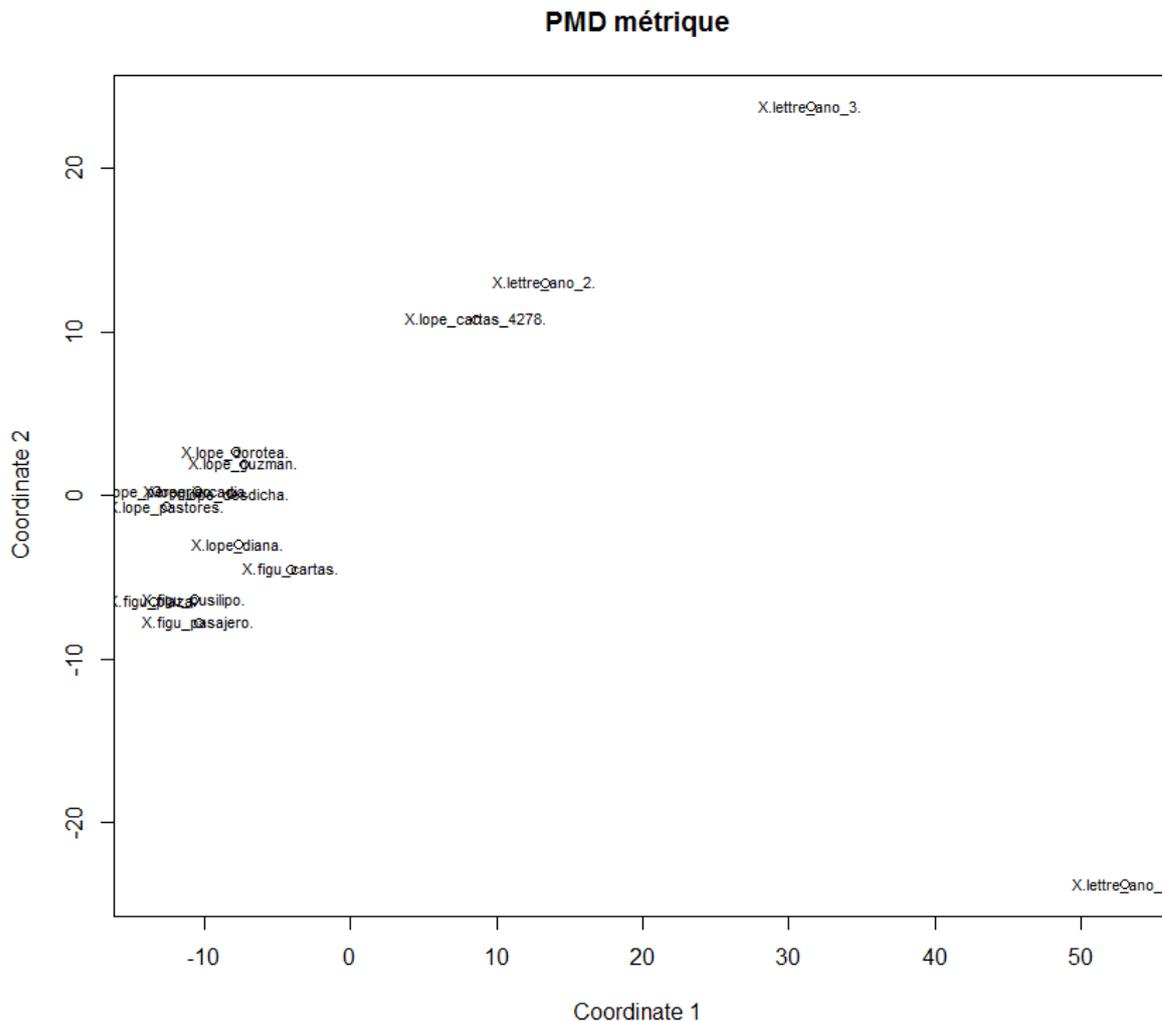


Figure 2 – Résultats graphiques du positionnement multidimensionnel

Le premier axe, horizontal, semble encore une fois repérer la différence générique entre lettres et fiction (sauf pour la correspondance de Suárez de Figueroa, ce qui ne laisse pas de troubler). Le second, vertical, est plus difficilement interprétable puisqu'il sépare nettement la première lettre anonyme (en bas à droite) des deux autres. Mise à part cette bizarrerie, l'opposition haut/bas semble recouper la différence auctoriale, avec les textes de Lope situés en haut à partir de l'ordonnée -2, et ceux de Suárez de Figueroa en dessous. En revanche, l'attraction entre la seconde lettre anonyme et la correspondance de Lope sur tableau lexical complet pondéré semble récurrente.

Analyses avec Stylo R : clustering et arbre de consensus

Le troisième test mis en œuvre, non plus en lignes de code informatique sur R mais directement sur l'interface de Stylo⁸⁴, considère un grand nombre de classifications en variant les paramètres (avec un *culling* de 50 à 100 %)⁸⁵, les méthodes utilisées (*clustering*, *consensus tree*), les observables statistiques (100 MFW, 20 MFW, 50 MFW, 100 cooccurrents de mots, 100 trigrammes de mots) et les mesures de distance : Manhattan, Canberra, classique. La configuration la plus performante s'est révélée être le *clustering* sur Delta de Burrows appliqué aux 100 MFW, avec un *culling* de 50 à 100 %. On obtient ce dendrogramme :

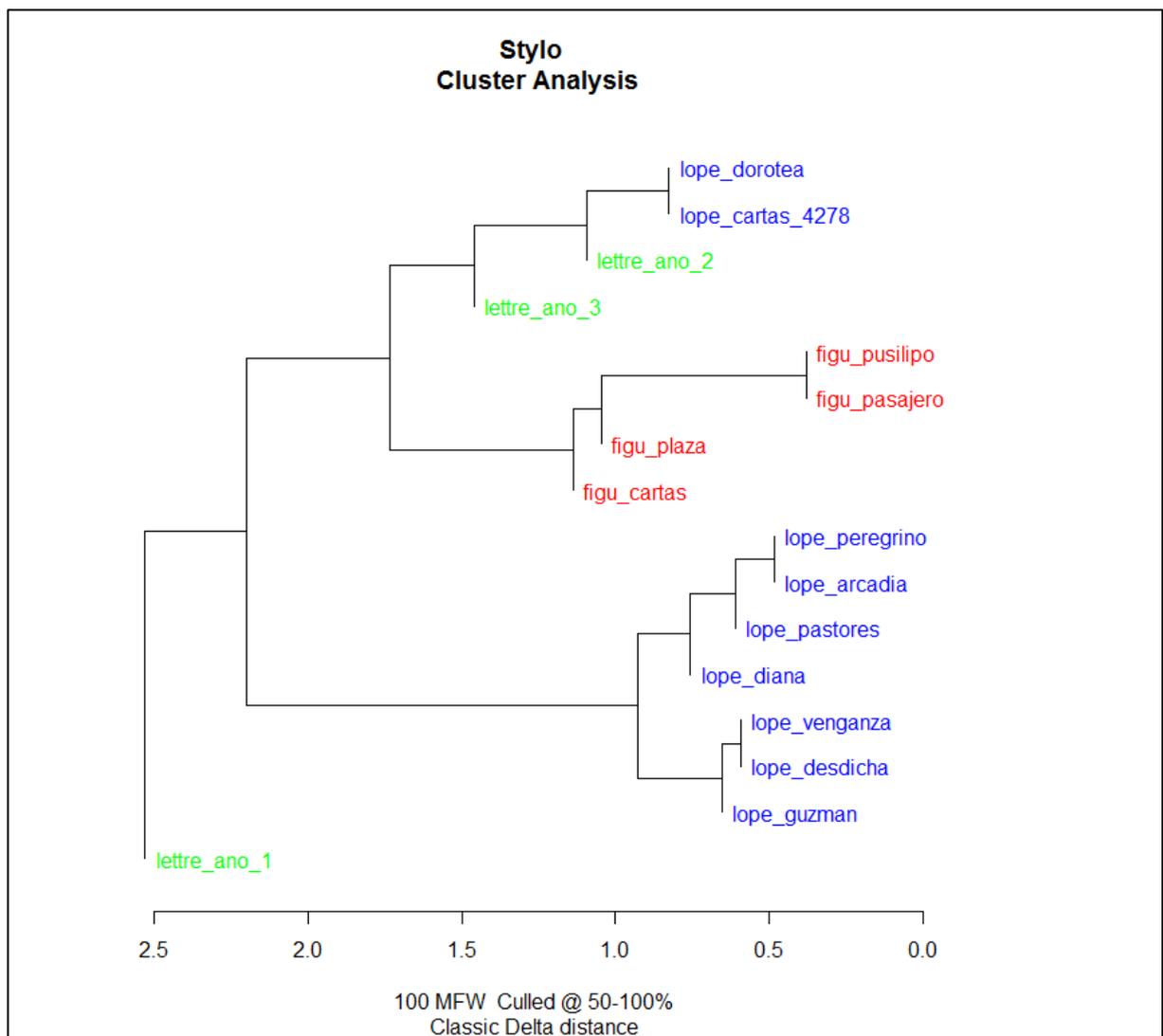


Figure 3 – Cluster Analysis sur Stylo R

Ce qui frappe au premier abord est la surprenante finesse de l'algorithme. Le programme a distingué le Lope épistolier et dialoguiste (lope_dorotea et lope_cartas) du Lope des nouvelles et des

⁸⁴ Voir Maciej Eder, Jan Rybicki et Mike Kestemont, *op. cit.*

⁸⁵ Un *culling* de 50 à 100 % désigne le fait que l'algorithme tourne uniquement sur les mots présents dans au moins la moitié des textes du corpus. L'intérêt est de réduire le biais de sélection, selon lequel des mots rares, et donc trop spécifiques à un texte en particulier, pourraient introduire un biais en faveur ou en défaveur du texte en question.

romans. Il a regroupé la plupart des *Novelas a Marcia Leonarda* ensemble (*Venganza, Desdicha, Guzmán*). Il distingue même les nouvelles du recueil *Circe* (*Venganza, Desdicha, Guzmán*) de celles de *La Filomena* (*Diana*). *Las fortunas de Diana*, classée avec les romans, se distingue en effet par un style plus travaillé ; son intrigue aussi fait plus penser à celle d'un roman que d'une nouvelle. L'algorithme distingue nettement, sur une branche à part, l'œuvre de Suárez de Figueroa, et sépare même les œuvres propres de l'auteur (*Pusílipo* et *Pasajero*, tous deux des miscellanées didactiques dialoguées) de sa traduction, qu'il classe à côté (la *Plaza universal de todas las ciencias*, traduction de l'italien *Piazza universale di tutte le professioni del mondo* de Tommaso Garzoni, 1585). Enfin, la branche du haut rapproche logiquement la *Dorotea* de la correspondance lopesque, sans doute en raison de l'emploi plus important des pronoms personnels de première et deuxième personne, typique du genre épistolaire et du dialogue. La lettre anonyme II est regroupée avec la correspondance de Lope, et, dans une moindre mesure, la lettre III. La première se trouve cependant rejetée au bas du graphique, comme si elle n'appartenait à aucun *cluster*. Sa taille extrêmement réduite ne permet pas de toutes façons de se fier aux résultats obtenus. L'analyse porte sur les 100 MFW du corpus, dont voici la liste :

de, y, que, la, en, el, a, los, con, por, se, las, no, su, del, es, como, lo, para, más, al, un, le, sus, si, tan, o, quien, una, ser, porque, sin, me, así, qué, bien, todo, son, pues, mas, mi, ni, entre, esta, todos, cuando, ya, cosas, yo, donde, este, fue, dijo, otros, había, dos, ha, era, esto, parte, haber, también, según, mismo, sobre, sólo, dice, él, sino, aquel, todas, te, cosa, vida, tal, tanto, aunque, dios, tiene, amor, hasta, tiempo, pero, ella, otro, muchos, menos, puede, otra, poco, otras, tierra, ojos, siempre, mayor, toda, después, hacer, dorotea, fin.

On a édité manuellement cette liste pour supprimer les quelques termes de contenu présents : *dios, amor, tiempo, tierra, ojos, dorotea*, remplacés par leurs successeurs en ordre de fréquence dans la liste : *siendo, cuanto, eso, fuera, hombre, hay*. La présence de mots propres à certaines œuvres de Lope (*dorotea*, en raison de la récurrence de l'apostrophe dans l'œuvre éponyme, *dios* et *tierra* dans les *Pastores de Belén*, roman sacré, *amor, tiempo*, et *ojos* dans la *Arcadia*, roman pastoral) aurait pu défavoriser la candidature de Lope en poussant à classer les lettres anonymes, qui ne contiennent pas ces mots ou très marginalement, comme éloignées de l'œuvre du Phénix. En fait, cela ne change pas les résultats, excepté que *figu_cartas* et *figu_plaza* s'intervertissent.

Pour vérifier la robustesse de ce résultat, nous avons varié la méthode (*consensus tree*) et les paramètres de l'analyse (ici le *culling*). Sur les 100 MFW et en utilisant la métrique de Burrows, en faisant varier le *culling* de 50 à 100 %, c'est-à-dire que l'on a demandé à l'algorithme de prendre en compte uniquement les mots présents dans la moitié des textes du corpus, puis de faire varier ce pourcentage par tranches (et analyses) successives jusqu'à un *culling* de 100 %, on obtient cet arbre :

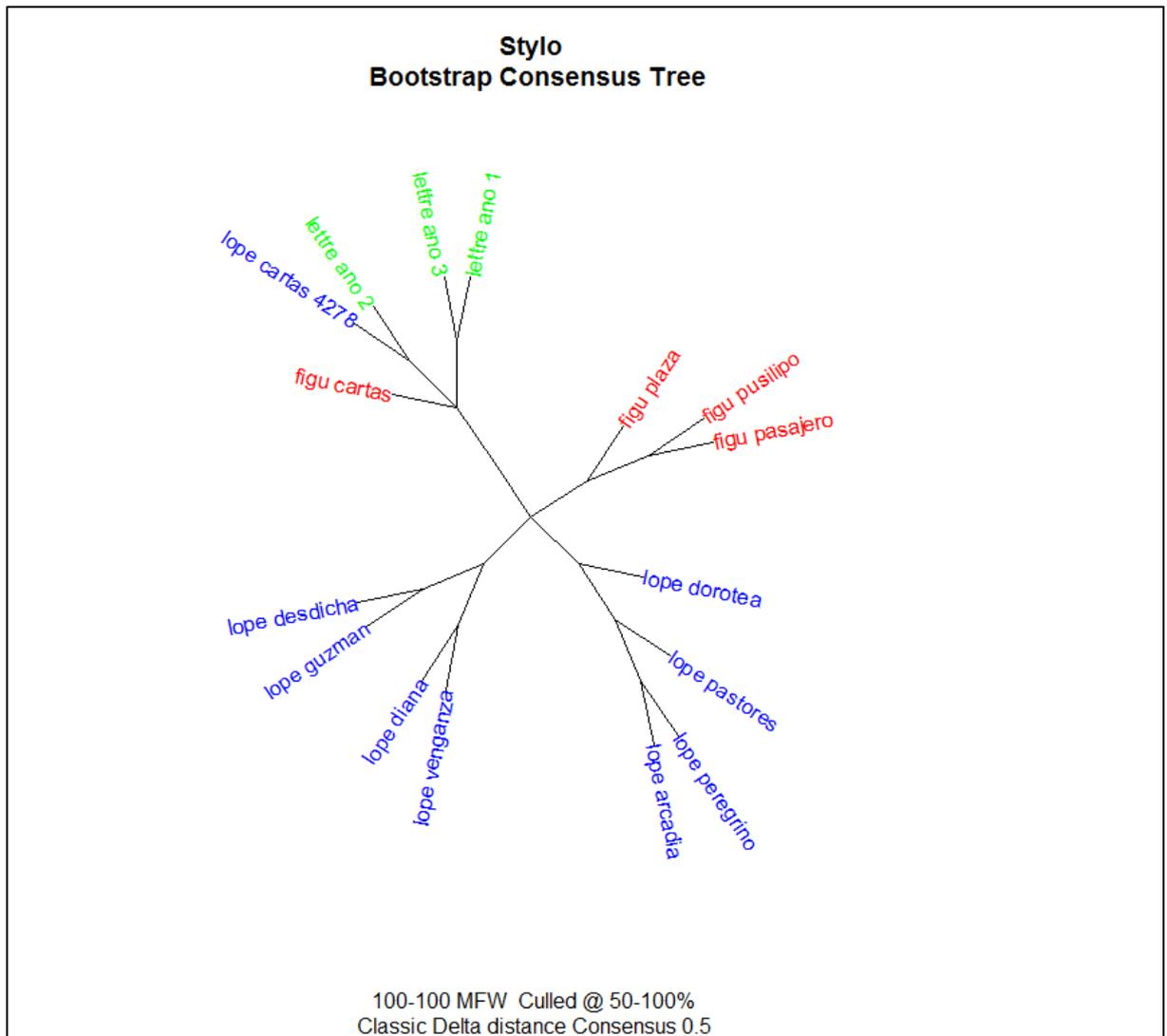


Figure 4 – Arbre de consensus sur Stylo R

L'arbre de consensus synthétise les résultats obtenus, en ne retenant que les *clusters* apparus dans au moins la moitié des résultats. Ici, l'analyse a correctement regroupé les romans de Lope ensemble, en bas à droite, l'œuvre de Suárez de Figueras en haut à droite, sauf les lettres qui posent problème ; les nouvelles de Lope en bas à gauche. Cette méthode semble très sensible au caractère épistolaire des textes, et classe, comme les méthodes testées précédemment, la lettre anonyme II avec les lettres de Lope, alors qu'elle sépare les lettres I et III dans une autre sous-branche, mais dans le même *cluster*.

Méthodes supervisées et résultats

Voyons maintenant les résultats obtenus en appliquant différentes méthodes de l'intelligence artificielle, dites aussi supervisées, pour réaliser la classification. Il s'agit de méthodes fondées sur l'établissement d'un corpus d'apprentissage : on crée un corpus d'entraînement, corpus de référence à partir duquel le programme apprend à reconnaître, de manière probabiliste, un auteur. On constitue un corpus d'entraînement (*primary_set*) contenant tous les textes du corpus, sauf les lettres anonymes (à

attribuer) et deux textes qui vont servir de test afin de mesurer la performance de l'algorithme (nous avons choisi le *Pusílipo* et le *Peregrino*) et un corpus de textes à attribuer (`secondary_set`) : les lettres, le *Pusílipo* et le *Peregrino*.

Les analyses supervisées sur R se lancent directement en lignes de commande. On a ici aussi testé tous les paramétrages possibles sur les différentes méthodes. La méthode des K plus proches voisins (*K nearest neighbors* ou KNN) consiste à étudier la distance d'un point non étiqueté (dans notre cas, les textes anonymes ou dont nous n'indiquons pas l'auteur) par rapport aux autres points, étiquetés (ceux qui portent en étiquette le nom d'auteur). On regarde à quel *cluster* appartiennent les points qui sont ses voisins les plus proches et on le met dans ce *cluster*. Les points, de même que tout à l'heure sur l'ACP, représentent les différents textes du corpus. On a d'abord testé les knn en mesurant les distances entre textes en prenant en compte les cent premiers bigrammes de mots (séquences de deux mots), les cent premiers 6grams de caractères ou groupements de six caractères consécutifs, les cent premiers 7-grams de caractères, les cent premiers 7-grams de mots, les 100 MFW, puis les mêmes traits avec la méthode des SVM, des NSC et de la classification naïve bayésienne. Les machines à vecteurs de support (*Support Vector Machines* ou SVM) ou séparateurs à vaste marge constituent une technique de classification visant à regrouper des individus linéairement en deux groupes tout en maximisant la distance entre les groupes. Cette approche serait l'une des plus performantes en attribution d'autorité⁸⁶. Quant au plus proche centroïde, il s'agit d'un algorithme itératif qui crée des *clusters* d'individus et minimise la somme des distances entre chaque individu et le centre de son *cluster*. Seule la classification naïve bayésienne échoue à attribuer le *Pusílipo* à Suárez de Figueroa et le *Peregrino* à Lope. La figure suivante comporte quelques-uns des résultats :

KNN sur bigrammes de mots	SVM sur bigrammes de mots	KNN sur 6grams de caractères
<pre> lettre_ano_1 --> figu lettre_ano_2 --> lope lettre_ano_3 --> lope 100 MFW , culled @ 40%, 2 of 2 (100%) </pre>	<pre> lettre_ano_1 --> lope lettre_ano_2 --> lope lettre_ano_3 --> lope 100 MFW , culled @ 40%, 2 of 2 (100%) </pre>	<pre> lettre_ano_1 --> figu lettre_ano_2 --> lope lettre_ano_3 --> lope 100 MFW , culled @ 40%, 2 of 2 (100%) </pre>
<pre> lettre_ano_1 --> figu lettre_ano_2 --> lope lettre_ano_3 --> lope 100 MFW , culled @ 60%, 2 of 2 (100%) </pre>	<pre> lettre_ano_1 --> lope lettre_ano_2 --> lope lettre_ano_3 --> lope 100 MFW , culled @ 60%, 2 of 2 (100%) </pre>	<pre> lettre_ano_1 --> figu lettre_ano_2 --> lope lettre_ano_3 --> lope 100 MFW , culled @ 60%, 2 of 2 (100%) </pre>
<pre> lettre_ano_1 --> figu lettre_ano_2 --> lope lettre_ano_3 --> lope 100 MFW , culled @ 80%, 2 of 2 (100%) </pre>	<pre> lettre_ano_1 --> figu lettre_ano_2 --> lope lettre_ano_3 --> lope 100 MFW , culled @ 80%, 2 of 2 (100%) </pre>	<pre> lettre_ano_1 --> figu lettre_ano_2 --> lope lettre_ano_3 --> lope 100 MFW , culled @ 80%, 2 of 2 (100%) </pre>
<pre> lettre_ano_1 --> lope lettre_ano_2 --> lope lettre_ano_3 --> lope 27 MFW , culled @ 100%, 2 of 2 (100%) General attributive success: 8 of 8 (100%) MFWs from 100 to 100 @ increment 100 Culling from 40 to 100 @ increment 20 Propouns deleted: FALSE </pre>	<pre> lettre_ano_1 --> lope lettre_ano_2 --> lope lettre_ano_3 --> lope 27 MFW , culled @ 100%, 2 of 2 (100%) General attributive success: 8 of 8 (100%) MFWs from 100 to 100 @ increment 100 Culling from 40 to 100 @ increment 20 </pre>	<pre> lettre_ano_1 --> lope lettre_ano_2 --> lope lettre_ano_3 --> lope 100 MFW , culled @ 100%, 2 of 2 (100%) General attributive success: 8 of 8 (100%) MFWs from 100 to 100 @ increment 100 Culling from 40 to 100 @ increment 20 </pre>

⁸⁶ Romain Brixtel, Charlotte Lecluze et Gaël Lejeune, « Attribution d'auteur : approche multilingue fondée sur les répétitions maximales », Actes de la 22^e conférence sur le traitement automatique des langues naturelles, 2015, p. 210.

Figure 5 – Échantillons de résultats obtenus en variant les méthodes et les observables

La lettre II est systématiquement attribuée à Lope ; ce résultat est robuste autant à la variation de méthodes, de paramètres, que de traits observés. Au regard de la récurrence de ce résultat à la fois par des tests non supervisés et supervisés, de l'attraction que subit la lettre II par rapport à la correspondance de Lope sur les graphiques, se trouvant toujours dans son voisinage, on peut établir avec une quasi-certitude que Lope de Vega est l'auteur de la *Carta de respuesta*.

De la stylométrie à l'exploration textométrique

Explorons maintenant le corpus en *close reading*, à l'aide de TXM, pour voir quels sont les traits qui concrètement rapprochent la lettre II de l'œuvre de Lope. En suivant la très juste interrogation de Hugh Craig : « *if you can tell authors apart, have you learned anything about them*⁸⁷? », on aimerait plaider pour relier stylométrie et textométrie, décloisonner les communautés, les faire dialoguer, et envisager les deux disciplines comme fondamentalement complémentaires : l'attribution d'autorité permet d'identifier des traits vraiment spécifiques des auteurs, et, de son côté, l'étude stylistique par la textométrie permet de revenir aux résultats de l'AA avec un regard éclairé par l'« indispensable retour au texte⁸⁸ ». On pourrait envisager les deux approches comme deux étapes essentielles du traitement d'un quelconque problème d'attribution d'autorité. Dans cette partie de l'analyse, sauf pour les spécificités sur TXM, nous avons regroupé les lettres anonymes entre elles, afin de mesurer la distance de nos deux auteurs cette fois-ci avec l'ensemble formé par les trois lettres.

Méthode

Nous avons extrait la table lexicale depuis TXM, pour travailler sur le fichier au format .csv sur Excel. Après avoir retenu les 100 MFW, nous avons tracé les courbes comparées de leurs fréquences relatives chez Lope, Suárez de Figueroa et dans les lettres anonymes. Nous obtenons les spectres lexicaux superposés, ce qui permet d'identifier des suremplois :

⁸⁷ Hugh Craig, « Authorial attribution and computational stylistics: if you can tell authors apart, have you learned anything about them? », *Literary and Linguistic Computing*, vol. 14, t. 1, avril 1999, p. 103. Nous traduisons: « Quand vous êtes capables de discriminer des auteurs, avez-vous appris quoi que ce soit sur eux ? »

⁸⁸ Bénédicte Pincemin, « Sémantique interprétative et textométrie – version abrégée », *Corpus*, num. 10, 2011, p. 259-269, paragr. 13.

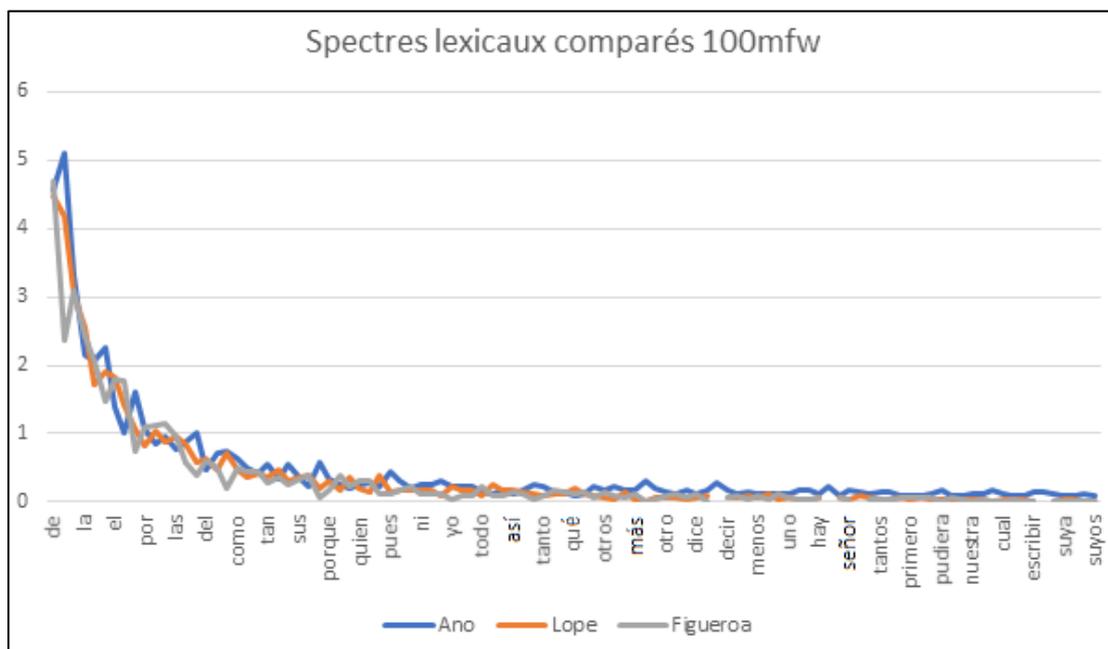


Figure 6 – Spectres lexicaux comparés des 100 MFW sur Excel

Pour chaque mot, on a ensuite calculé manuellement, sur Excel, la différence des fréquences moyennes d'un auteur à l'autre, pondérée par la taille des échantillons. On en a pris la valeur absolue, avant de calculer la moyenne arithmétique de ces valeurs. On exprime le résultat en pourcentage, qui correspond au pourcentage de variation de chaque auteur candidat aux lettres anonymes. La procédure correspond en fait à la « procédure delta » de John Burrows⁸⁹.

À partir de là, toujours sur Excel, on identifie les mots-outils les plus employés dans les lettres anonymes : il s'agit de *que, a, no, si, un*. *Que* présente dans les lettres une valeur de 5,09, proche de sa valeur chez Lope (4,18), et très éloignée de celle de Suárez de Figueroa (2,35). C'est la même chose pour la préposition *a*, avec une valeur de 2,27 dans les lettres, 1,92 chez Lope et 1,48 chez Suárez de Figueroa. Il en va de même pour *no* (1,60 dans les lettres et 1,06 chez Lope contre 0,26 chez Suárez de Figueroa) et *si* (respectivement 0,55 et 0,31 contre 0,26), avec la notable exception de *un* (avec des valeurs de 0,35, 0,46 et 0,35), dont l'emploi dans les lettres est plus proche de l'auteur de la *Plaza*.

	Lettres anonymes	Lope	Suárez de Figueroa
Que	5,09	4,18	2,35
A	2,27	1,92	1,48
No	1,60	1,06	0,73
Si	0,55	0,31	0,26
Un	0,35	0,46	0,35

Tableau 2 – Fréquences comparées des premiers MFW

⁸⁹ À ne pas confondre avec le « delta de Burrows », mesure de distance.

Sur l'ensemble du spectre lexical, les fréquences d'emploi dans les lettres présentent une variation de 47 % par rapport à Lope de Vega, et 50 % par rapport à Suárez de Figueroa. Les proportions d'emplois des 20 MFW, et en particulier des mots suremployés par les lettres anonymes, sont donc très proches entre celles-ci et Lope.

Étant donnée cette proximité de Lope aux lettres, on a voulu comparer les spécificités des 20 mots les plus fréquents des différents textes en utilisant le modèle hypergéométrique de Pierre Lafon, implémenté sur TXM.

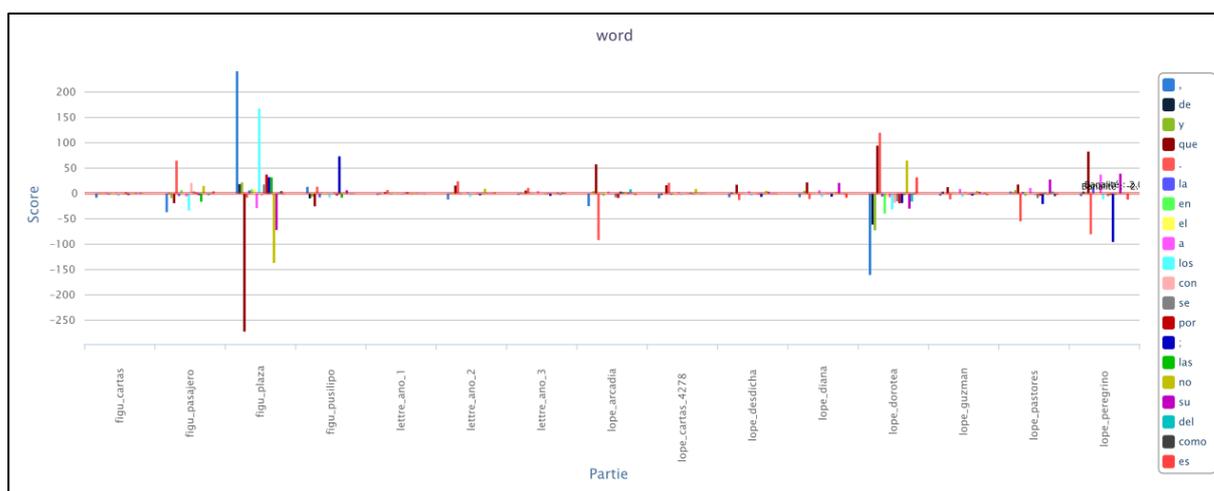


Figure 7 – Spécificités des 20 MFW sur TXM

Un zoom permet de mieux apprécier l'extrême similarité du profil de spécificités entre la lettre II et la correspondance de Lope :

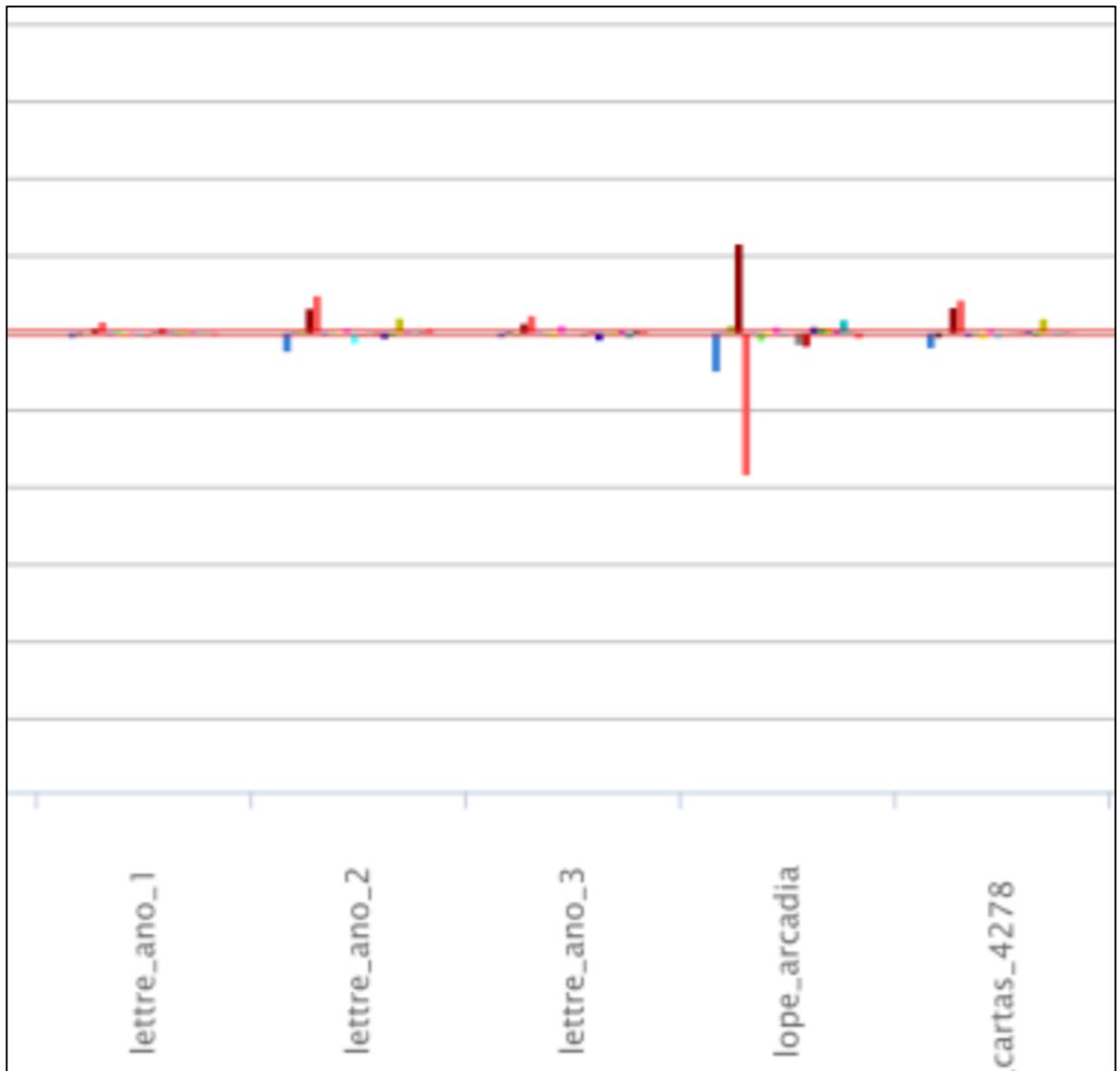


Figure 8– Zoom sur la figure précédente

Bien entendu, une simple ressemblance graphique ne suffit pas à tirer de conclusion pour l'autorité de la lettre II. Cependant, un coup d'œil aux valeurs associées aux 20 MFW dans les deux cas révèle une proximité édifiante entre les lettres et Lope.

Exploration des lettres en *close reading* à partir des résultats quantitatifs

Après nous être intéressée aux marqueurs inconscients du style que sont les mots-outils, il nous semble pertinent, dans l'optique d'un travail en humanités numériques *littéraires*, d'observer manuellement, par une analyse traditionnelle, les effets rhétoriques associés aux mots-outils détectés comme quantitativement saillants chez nos auteurs. Nous nous en tiendrons ici à Lope de Vega faute de

place, ainsi que pour aller dans le sens des résultats probabilistes obtenus en stylistique computationnelle.

L'étude qualitative est envisagée ici comme une étape, chronologiquement seconde, mais logiquement importante, de l'analyse stylo-métrique⁹⁰. Il s'agira donc dans cette partie d'étudier la partie plus consciente du style, autrement dit la manière dont s'insèrent, à l'échelle globale de la phrase, les marqueurs inconscients détectés par l'analyse quantitative. D'ailleurs, même s'il est hors de propos de discuter ici cette question, il n'est pas dit que l'échelle de la phrase constitue le pré carré d'un maniement conscient de la langue car, d'après Barthes que nous suivons ici, c'est le style dans son entier, non seulement pris du point de vue microscopique, mais aussi macroscopique, qui constitue le produit incontrôlé du corps et de la psyché de l'auteur.

Plus précisément, on souhaite ici étudier l'emploi de la conjonction de subordination et de l'adverbe d'hypothèse dans les lettres anonymes, et en particulier dans les deux dernières, afin de montrer que le style de ces épîtres est très proche de celui de Lope. On s'emploiera ainsi à identifier leurs stratégies rhétoriques partagées, afin de démontrer qualitativement cette similarité : non seulement les trois textes se rejoignent entre eux, mais ils ressemblent de manière troublante aux lettres de Lope, de par leur agressivité et leurs balancements permanents entre louanges feintes et attaques explicites.

Nous ne nous livrerons pas ici à une étude exhaustive de ces postures, mais nous contenterons de donner quelques pistes pour une future étude comparée des lettres anonymes et des épîtres de Lope. Les exemples retenus sont obtenus par un retour au texte à partir de requêtes sur l'index de TXM portant sur « que » et « si », assortis d'un ou deux exemples tirés de notre lecture personnelle du corpus.

Des lettres courtoises et apparemment bienveillantes

Avant d'esquisser quelques éléments de comparaison entre le style et le contenu des lettres anonymes et celles de Lope, il convient de vérifier manuellement la proximité du style des trois lettres anonymes entre elles. Après examen des instances liées à l'adverbe d'hypothèse et à la conjonction de subordination, il apparaît qu'elles présentent effectivement à la fois un style et une attitude proche, celle d'un ami prêtant conseil à un jeune poète encore inexpérimenté et dont les erreurs, bien naturelles, ne constituent qu'une étape inévitable et éphémère de sa carrière poétique.

Ce passage de la première lettre est un bon exemple de cette attitude de conseil empreinte de louanges valant *captationes benevolentiae* :

Y caso (no lo permita Dios) que V. m. por mostrar su agudeza quiera defender que merece alabanza por inventor de dificultar la construcción del romance, no se deje caer V. m. en esta tentación, ya que tiene tantos ejemplos de mil ingenios altivos que se han despeñado por no reconocer su primero disparate⁹¹.

⁹⁰ Nous avons essayé de décrire et d'appliquer ce double temps de l'analyse stylo-métrique dans un travail sur les séquences grammaticales récurrentes chez Góngora : Marie-Églantine Lescasse, *art. cit.*, p. 115-116.

⁹¹ Antonio Carreira, *art. cit.*, p. 161. Nous traduisons : « Et si à tout hasard (mais que Dieu ne le permette pas !) pour montrer votre ingéniosité, vous vouliez défendre l'idée selon laquelle vous méritez d'être loué pour avoir découvert une manière de

L'habileté de cette période consiste en l'alternance entre une attitude qui ménage son lecteur et des remarques quelque peu incisives, passant de la reconnaissance de l'ingéniosité du poète (« *su agudeza* ») et son appartenance à la catégorie des « esprits fiers » (« *ingenios altivos* ») à l'idée de complexification délétère de la syntaxe castillane et de chute prochaine prédite au poète, s'il se montre incapable de reconnaître ses erreurs.

L'auteur des lettres anonymes déguise sa critique en louange et en conseil, et a l'intelligence de présenter son destinataire comme victime d'une erreur momentanée de jugement. Les lettres demeurent ainsi toujours extrêmement courtoises et partagent toutes le thème de l'erreur, celle de se glorifier d'avoir introduit un nouveau style en poésie castillane, excusable, car due à un aveuglement passager, que l'auteur des lettres pourrait aider à corriger. Elles ne mettent pas sur le compte de Góngora une incompétence dangereuse mais des malentendus, rectifiables pour qui sait écouter l'aimable conseil d'un ami. C'est également le cas dans la lettre III où la querelle avec Lope est présentée comme une suite de quiproquos. Dans la deuxième lettre, le conseil de laisser mûrir son œuvre participe également de cette attitude bienveillante, destinée à orienter un jeune poète pas encore – vraisemblablement – débarqué à la cour⁹².

De ce point de vue, le point de comparaison le plus opportun sont les lettres publiées, quelques années plus tard, par Lope dans deux de ses livres, *La Filomena* (1621) et *La Circe* (1624), et où il critique ouvertement, mais de manière courtoise et prudente, la nouvelle manière de Góngora (qu'il s'abstient pourtant de nommer) et surtout sa diffusion chez d'autres poètes qui, malencontreusement, se sont mis à l'imiter. Une différence massive et évidente est que la critique est à la troisième personne, il n'y a pas de prise à parti directe, d'interpellation.

Dans ces lettres, le Phénix se plaît à étourdir son lecteur par la juxtaposition de longues subordonnées au sens souvent contradictoire, destinées à éveiller la perplexité et à dissimuler des critiques sous des dehors ambivalents. C'est le cas dans ce passage des épîtres de *La Filomena* :

Mas no contento con haber hallado en aquella blandura y suavidad el último grado de la fama, quiso [...] enriquecer el arte y aun la lengua con tales exornaciones (f^o 193 r^o) y figuras, cuales nunca fueron imaginadas ni hasta su tiempo vistas, aunque algo asombradas de un poeta en idioma toscano que por ser de nación genovés no alcanzó el verdadero dialecto de aquella lengua, donde hay tantas insignes obras inteligibles a la primera vista de los hombres doctos y aun casi de los ignorantes. Bien consiguió este caballero lo que intentó, a mi juicio, si aquello era lo que intentaba: la dificultad está en el recibirlo, de que han nacido tantas, que dudo que cesen, si la causa no cesa⁹³.

rendre la construction de la langue vulgaire plus difficile, ne cédez pas à cette tentation, car il y a tant d'exemples de mille esprits fiers qui ont chuté pour n'avoir pas reconnu leur ineptie première ».

⁹² « *Si V. m. f. como lo dice, fuera observante de los preceptos de Horacio, dejara reposar sus obras.* » Emilio Orozco Díaz, *op. cit.*, p. 240. Nous traduisons : « Si, comme vous le dites, vous observiez vraiment les préceptes d'Horace, vous laisseriez votre œuvre au repos ».

⁹³ Félix Lope de Vega, *Epístolas de La Filomena, op. cit.*, f^o 192 v^o-193 r^o. Nous traduisons : « Mais, ne se satisfaisant pas d'avoir trouvé dans ce style doux et amène le dernier degré de la célébrité, il voulut [...] enrichir l'art et même la langue avec des ornements et des figures telles qu'elles n'avaient jamais été imaginées ni vues auparavant, bien qu'un peu étonnantes de la part d'un poète de langue toscane, qui, parce que Génois d'origine, n'avait jamais atteint au maniement parfait de cette langue, dans laquelle il existe tant d'œuvres insignes, intelligibles au premier coup d'œil aux hommes doctes et même presque aux

Le superlatif « *último* » et l'adjectif d'intensité « *tales* », à valeur laudative, font place à une concessive ambiguë (« *aunque algo asombradas de un poeta...* »). De manière assez perfide, on trouve à Góngora un précurseur, et comme une préfiguration de sa manière, ce qui lui ôte l'avantage de l'invention absolue (« *cosas nunca imaginadas ni hasta su tiempo vistas* »), avec un poète en langue toscane et d'origine génoise (il s'agit de Gabriello Chiabrera) incapable d'écrire dans une langue pure et correcte, tandis que le génitif sujet « *de doctos o ignorantes* » réactualise de manière discrète la critique souvent faite à Góngora de plaire à la plèbe ignorante (*vulgo*), émerveillée par ses acrobaties langagières auxquelles elle ne comprend rien. Le polyptote « *cesen* »/ « *cesa* » suggère, via l'identité verbale, la nécessité d'arrêter d'écrire : seule la démission du poète pourra mettre un terme aux difficultés de la réception, renoncement salutaire selon Lope pour le lecteur, empêtré dans l'obscurité syntaxique et lexicale du Cordouan.

On retrouve dans ce fragment plusieurs caractéristiques de la citation précédente issue de la lettre II, jusque dans la construction même, depuis l'idée d'une gloire littéraire déjà acquise conformant la protase (« *fama* » / « *agudeza* ») à celle d'une volonté apparaissant comme capricieuse et arbitraire, de modifier la syntaxe et l'ornement castillan jusqu'à en aliéner la langue. Les deux fragments utilisent le verbe « vouloir » (« *y caso que V.m. quiera defender* » / « *quiso [...] enriquecer el arte y aun la lengua con tales exornaciones* »), des arguments proches, faisant se suivre une proposition encomiastique et l'idée d'une conséquence linguistique inédite et néfaste.

Conclusions

Au vu de la récurrence de l'attribution de la lettre II à Lope par Stylo, malgré la variation des approches (supervisées avec corpus d'apprentissage / non supervisées), des métriques (Delta de Burrows, de Manhattan, de Canberra), des méthodes (ACP, MDS, KNN, SVM, NSC, classification naïve bayésienne), des paramètres (*culling* de 20 à 100 %) et des observables choisis (100 MFW, 20 MFW, bigrammes de mots, bigrammes de caractères, 7-grams de mots, 7-grams de caractères), au vu aussi de la proximité des spectres lexicaux obtenus par une analyse manuelle sur Excel après extraction des tables sur TXM, des résultats de l'analyse des spécificités sur les 20 MFW sur TXM, au vu enfin des similitudes de postures et de stratégies rhétoriques entre la lettre II et les lettres de Lope que nous a permis d'identifier l'analyse qualitative, nous pouvons attribuer avec une probabilité très élevée la seconde lettre anonyme à Lope.

Les conclusions sont beaucoup moins certaines pour les lettres I et III, entre autres en raison de leur taille réduite, inférieure à la limite recommandée en attribution d'autorité.

ignorants. Ce gentilhomme a bien réussi ce qu'il s'était proposé, à mon avis, si c'était bien cela qu'il visait : la difficulté réside à présent dans sa réception, réception à partir de laquelle naquirent tant de difficultés que je doute que celles-ci ne cessent, si la cause ne cesse pas. »

(Pas de bibliographie en fin d'article)