



HAL
open science

Louise Labé : une créature de papier?

Florian Cafiero, Jean-Baptiste Camps, Simon Gabay

► To cite this version:

Florian Cafiero, Jean-Baptiste Camps, Simon Gabay. Louise Labé : une créature de papier? : Sur l'attribution de ses poèmes à Olivier de Magny. *Humanistica 2023*, Association francophone des humanités numériques, Jun 2023, Genève, Suisse. hal-04090284

HAL Id: hal-04090284

<https://hal.science/hal-04090284>

Submitted on 5 May 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.



Distributed under a Creative Commons Attribution 4.0 International License

Louise Labé: une créature de papier? Sur l'attribution de ses poèmes à Olivier de Magny

Florian Cafiero

Sciences Po Paris

prenom.nom@sciencespo.fr

Jean-Baptiste Camps

Ecole nationale des chartes | PSL

prenom.nom@chartes.psl.eu

Simon Gabay

Université de Genève

prenom.nom@unige.ch

Résumé

Depuis plusieurs années, l'idée que Louise Labé ne soit que le prête-nom de quelques poètes lyonnais s'est imposée comme une hypothèse plausible parmi les spécialistes du XVI^e s. Dans ce débat, aucun protagoniste n'a pour l'instant mobilisé les approches computationnelles. Le présent travail se propose donc de relancer la discussion au moyen de la stylométrie, et remet en doute l'attribution de l'œuvre en vers de Labé à Magny.

Parmi les plus important-e-s auteur.trice.s des premiers siècles de la littérature française, on compte Louise Labé et ses célèbres *Euvres* (Labbé, 1555). Les points médians ont ici toute leur importance : est-iel une femme ou le prête-nom de quelques poètes talentueux ? D'après l'essai *Louise Labé, une créature de papier* de Mireille Huchon paru en 2005 (Huchon, 2005), la question se pose, et la réponse pencherait sérieusement vers la seconde hypothèse.

La thèse peut être résumée simplement : Louise Labbé, courtisane lyonnaise du XVI^e s. à la vie tumultueuse, ne serait pas Louise Labé, nom de plume derrière lequel se cache un petit groupe d'auteurs lyonnais. Si cette hypothèse a été reçue positivement par certains chercheurs (Fumaroli, 2006), d'autres ont émis quelques sérieuses réserves à propos de la théorie (Martin, 2006; Buron, 2006) ou de faits allégués pour la soutenir (Jourde, 2017) sans cependant qu'aucun parti n'arrive à définitivement convaincre l'autre.

La parution des *Œuvres complètes* de « Louise Labé » par la même M. Huchon (Labé, 2021) relance le débat, mobilisant toujours plus d'arguments tirés de la critique externe (témoignages d'époque, histoire littéraire. . .) et interne (emprunts, thèmes. . .). À ce jour, personne n'a cependant tiré partie des dernières avancées de la stylométrie. Cette dernière s'est pourtant imposée comme essentielle dans les problèmes d'*authorship attribution*. Le présent article se propose donc d'offrir, dans

la lignée de précédentes expériences (Cafiero and Camps, 2019; Gabay, 2021), un éclairage nouveau sur ce débat désormais ancien.

1 Description du problème

L'œuvre complète de Louise Labé est de taille très réduite. Elle se résume à une épître liminaire, un *Debat de Folie et d'Amour*, vingt-quatre sonnets (dont un en italien) et trois élégies, qui nous sont parvenus dans un imprimé issu des presses de Jean de Tournes (1504-1564). En plus de ces textes, on trouve dans ce recueil des *Escriz de divers Poètes à la louenge de Louïze Labé Lionnoïze* qui sont l'œuvre des plus grands auteurs de l'époque : Jean-Antoine de Baïf (1532-1589), Pontus de Tyard (1521-1605), Maurice Scève (c. 1505-1569). . . D'après M. Huchon, quelques-uns de ces auteurs se cacheraient derrière le personnage fictif de Louise Labé : l'épître des *Euvres* de 1555 serait de Claude de Taillemont (1504 -1558), le *Débat* de Maurice Scève (c. 1505-1569), et les pièces poétiques d'Olivier de Magny (c. 1529-c. 1561).

Ces trois auteurs sont tous membres de la célèbre école lyonnaise. Maurice Scève, en est sans doute l'un des principaux représentants avec les célèbres dizains de *Délie* (Scève, 1544) et quelques plus longs poèmes comme *La Saulsaye* (Scève, 1547) ou *Microcosme* (Scève, 1562). Claude de Taillemont, bien moins célèbre, est l'auteur de *La Tricarite*, composée de douzains suivis de quelques chants (Taillemont, 1556). Olivier de Magny enfin, parfois présenté comme un amant de Louise, a essentiellement composé des sonnets et des odes, parus dans ses *Amours* (Magny, 1553), ses *Soupirs* (Magny, 1557) et ses *Odes* (Magny, 1559).

2 Corpus de travail

2.1 Délimitation du corpus

La présente expérience se concentre sur l'étude du corpus en vers – le *Débat* comme l'épître limi-

naire sont donc exclus – pour un total d’un peu plus de 4 000 tokens. On sait en effet que le genre (comédie, tragédie. . .) comme la forme (vers ou prose) des textes font partie des nombreux paramètres à contrôler dans le cas de l’analyse stylométrique, dans la mesure où ils ont leur propre signal (Chaudhuri et al., 2019), et peuvent donc brouiller celui de l’auteur. Concernant la taille du corpus, on sait aussi que 4 000 tokens est suffisant, un minimum de 2 000 (Eder, 2017), voire 1 000 (Hirst and Feiguina, 2007) étant préconisé pour obtenir des résultats fiables.

La stylométrie étant par essence comparative, il est nécessaire de confronter les écrits de plusieurs auteurs : l’auteur·trice contesté·e, celui ou celle à qui les documents serait réattribués, et des auteurs de contrôle ne posant aucun problème d’attribution afin de garantir la fiabilité des résultats.

À cette fin, ont été choisis des poèmes :

- deux auteurs putatifs :
 - de *Délie* (1544) de Maurice Scève (c. 1505-c. 1569);
 - des *Amours* (1553) d’Olivier de Magny (c 1529-c. 1561);
- deux autres auteurs des cercles lyonnais :
 - du *Repos de plus grand travail* (1550) de Guillaume des Autels (1529-c. 1599);
 - des *Erreurs amoureuses* (1549) de Pontus de Tyard (1521-1605);
- quatre autrices :
 - des *Rithmes et poésies* (1546) de Pernelle du Guillet (c. 1518-1545);
 - des *Emblèmes ou devises chrestiennes* (1571) de Georgette de Montenay (c. 1540-1581);
 - des *Secondes Œuvres* (1583) de Catherine (1542-1587) et Madeleine (1520-1587) Des Roches.

Parmi les textes de ces auteur·ice·s, on trouve de multiples formes poétiques : sonnets, épigrammes, chants, dizains. . .

2.2 Sélection des *features*

Le système graphique varie considérablement (relativement à l’époque) d’un imprimé à un autre, créant un possible signal lors de l’analyse stylométrique. Afin de pallier ce problème, les textes ont été lemmatisés avec un modèle dédié au français de l’époque moderne en orthographe d’origine (Gabay et al., 2020b,a), et cette lemmatisation a été partiellement corrigée en ce qui concerne les mots-

outils, via l’interface Pyrrha (Clérice et al., 2022). Les œuvres de chaque auteur ont ensuite été découpées en échantillons de 1000 mots consécutifs. Les mots-outils en ont ensuite été extraits, en raison de leur capacité à caractériser l’idiolecte de chaque auteur (Kestemont, 2014), et leur fréquence calculée. Plus précisément, nous avons retenu les indéfinis, pronoms, déterminants, prépositions, conjonctions, et adverbes les plus courants. Nous avons toutefois exclu les interjections, ainsi que les pronoms personnels et possessifs, fortement sujets aux biais de genre littéraire, de même que les adverbes en *-ment* construits sur des adjectifs ou des substantifs, qui sont trop riches d’information sémantique pour constituer proprement des mot vides.

Le choix des lemmes de mots-outils, à l’exclusion d’autres caractéristiques performantes en attribution (tels que les affixes) se justifie ici par le besoin de la normalisation graphique par lemmatisation, mais d’autres caractéristiques pourraient faire l’objet de futures enquêtes.

En raison de la petite taille des échantillons, et pour garantir la fiabilité des analyses, nous ne retenons en dernier lieu que les variables pour lesquelles, d’un point de vue statistique, la fréquence est suffisante pour se situer, avec 90% de certitude, à plus ou moins deux écarts-types de la fréquence théorique, suivant une procédure proposée par H. Moisl et inspirée des méthodes d’enquête (Moisl, 2011; Cafiero and Camps, 2019).

3 Protocole expérimental

En raison de l’absence d’œuvres sûrement attribuées à Louise Labé, il ne nous est pour le moment possible que de traiter le problème de l’attribution des *Euvres* de Louise sous deux angles :

1. L’un des candidats les plus fréquemment proposés est-il régulièrement plus proche stylométriquement des œuvres de Louise Labé (et donc susceptible de les avoir écrites) ?
2. Existe-t-il un signal autorial indépendant dans les œuvres attribuées à Louise Labé, et raisonnablement distinct de celui d’autres auteurs et autrices de son temps ?

3.1 La méthode des imposteurs

Pour aborder la première problématique, nous choisissons une méthode de vérification d’autorité bien connue, celle dite des « imposteurs » ou *General Imposters* (Seidman, 2013; Koppel and Winter,

2014; Kestemont et al., 2016). Plutôt que d’étudier directement les distances entre deux textes, elle consiste à comparer le texte à attribuer x à l’ensemble des textes disponibles pour un auteur, $T = \{t_1, \dots, t_n\}$, ainsi qu’à chacun des textes d’un ensemble de n auteurs de contrôle (*distractors*, les ‘imposteurs’) $I = \{i_1, \dots, i_n\}$. Il s’agit de voir si x est régulièrement plus similaire à un texte de T qu’à un texte de I .

Pour ce faire, un échantillon des variables est prélevé à chaque itération dans celles décrivant le texte à attribuer, ainsi que celles des textes du candidat examiné et d’une sélection aléatoire des textes des auteurs de contrôle. Il s’agit ensuite de déterminer quel texte, parmi ceux du candidat et des contrôles, est le plus proche de celui à attribuer, selon une métrique à définir. On répète ensuite cette opération pour un nombre donné d’itérations, et on calcule la proportion dans laquelle le texte a été plus proche (*i.e.*, si la valeur est 1, le texte à attribuer a toujours été plus proche d’un texte du candidat; si la valeur est 0, il n’a jamais été plus proche d’un texte du candidat).

Si la proportion est inférieure à un seuil P_1 , on peut conclure à une différence significative (le candidat ne semble pas être l’auteur); si elle est supérieure à un seuil P_2 , à une proximité significative (le candidat pourrait être l’auteur). Pour déterminer ces deux seuils, on peut suivre une procédure d’optimisation : différents seuils sont testés sur les données connues (les textes des candidats et des contrôles), jusqu’à trouver ceux qui minimisent l’erreur d’attribution. Dans une situation idéale, les valeurs de P_1 et P_2 sont identiques (il n’y a pas de zone grise entre les deux), mais ce n’est pas nécessairement le cas.

Pour tirer le meilleur parti d’un corpus somme toute limité, nous paramétrons en choisissant aléatoirement 50% des variables et 75% des textes de contrôle à chacune des 500 itérations. Nous choisissons le delta de Burrows comme métrique (Burrows, 2002), et nous utilisons l’implémentation des imposteurs proposée par le module Stylo (Eder et al., 2016). La procédure d’optimisation a déterminé les seuils suivants : $P_1 = 0.27$ et $P_2 = 0.30$.

3.2 L’Analyse factorielle des correspondances

La méthode des imposteurs présente toutefois des risques : dans la mesure où elle se fonde uniquement sur le texte le plus proche, elle est très sensible à des rapprochements ponctuels dans les

données. Pour limiter ce problème, nous choisissons de recourir dans un second temps à une méthode non supervisée de visualisation, l’analyse factorielle des correspondances ou AFC (Benzécri, 1980), nous permettant de visualiser en deux dimensions, les nuages de points des textes des auteurs, pris trois par trois pour pouvoir être opposés en deux dimensions seulement.

Nous utilisons l’implémentation de l’AFC par le module FactoMineR (Lê et al., 2008).

4 Résultats

4.1 Expérience 1 : méthode des imposteurs

	Score imposteur
Des Roches	0.03
Guillet	0.04
Des Autels	0.08
Scève	0.11
Montenay	0.13
Magny	0.21
P_1	0.27
P_2	0.29
Tyard	0.70

TABLEAU 1 – Résultats de la procédure des imposteurs

Les résultats de la procédure des imposteurs (tabl. 1) ne donnent aucun candidat dans la *zone grise* entre P_1 et P_2 . Tous les candidats, y compris Magny, sont éliminés, à l’exception d’un unique candidat qui se distingue.

Un seul auteur se distingue parmi les candidats étudiés : Pontus de Tyard. Notons que cela ne permet pas de conclure qu’il ait écrit les œuvres de Louise Labé, mais uniquement qu’il est, de très loin, le meilleur candidat *si et seulement si l’auteur se trouve parmi ceux testés*. Pour préciser les choses, nous recourons à une deuxième analyse.

4.2 Expérience 2 : AFC

Pour préciser ce résultat, nous avons recours à une méthode de réduction de la dimensionalité, dont l’usage est ancien et répandu en stylométrie (Benzécri and Meïmaris, 1996), et parmi celles disponibles, nous retenons l’analyse factorielle des correspondances. Nous y confrontons les œuvres signées du nom de Louise Labé à celle de Pontus de Tyard et à celle du moins mauvais des candidats éliminés, Magny. Si les variations idiolectales sont les plus fortes au sein du corpus choisi, chacun des

deux axes du plan factoriel devrait délimiter un espace dévolu aux œuvres d'un auteur, opposées à celle d'un autre.

Le résultat de cette analyse est plutôt tranché (fig. 1) : si les textes signés de la main de Louise Labé sont effectivement plus distants (et opposés par un axe à la significativité plus grande) de ceux de Magny que de ceux de Pontus de Tyard, en revanche, ils semblent toutefois former un nuage distinct. En l'état, en l'absence de recouvrement partiel ou total des points des textes de Louise Labé et de ceux d'un autre auteur, rien ne permet en l'état de conclure que les textes signés de Louise Labé aient été écrits par quelqu'un d'autre. Notons toutefois que cela ne suffit pas tout à fait non plus à infirmer que le 'style Labé' soit celui d'un individu tiers ou le résultat d'une collaboration.

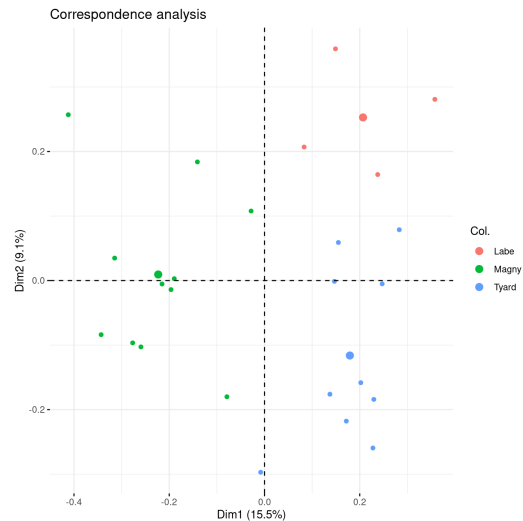
5 Conclusion et perspectives de recherche

Si les méthodes computationnelles n'ont pas vocation à remplacer le commentaire du texte et l'histoire du livre, il convient de rappeler leur importance croissante, voire cruciale, pour certains débats comme l'attribution d'auteur. De nombreux autres traitements seront bien sûr nécessaires pour approfondir la question complexe posés par ces textes singuliers. Des méthodes de profilage (Cafiero and Puren, 2022; Cortelazzo et al., 2018) pourraient permettre d'en affirmer plus quant au genre de la personne ayant écrit ces textes. Une étude par stylométrie roulante (Eder, 2016) pourrait également permettre de détecter d'éventuels changements de main au sein du corpus (Cafiero and Camps, 2021; Cafiero and Camps, 2022), et de confirmer que toutes les œuvres de Louise Labé sont bien écrites par une seule et même plume.

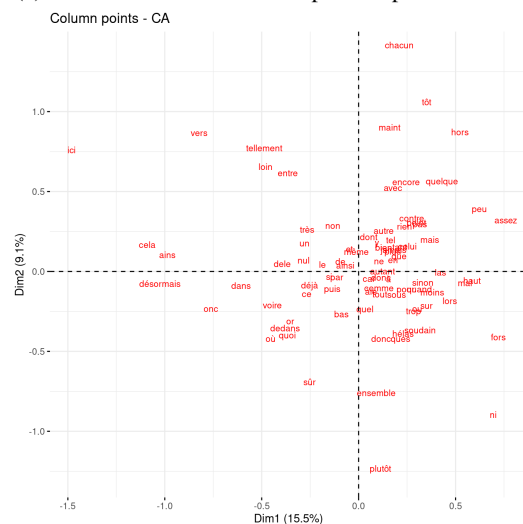
En regard, toutefois, à notre question initiale, nos résultats actuels semblent éliminer Magny comme (unique) auteur des œuvres de Louise Labé, faire de Pontus de Tyard un éventuel meilleur candidat, tout en semblant montrer que les œuvres de Louise Labé possèdent un signal propre et une forme d'unité. La petite taille actuelle du corpus interdit toutefois encore toute conclusion ferme.

References

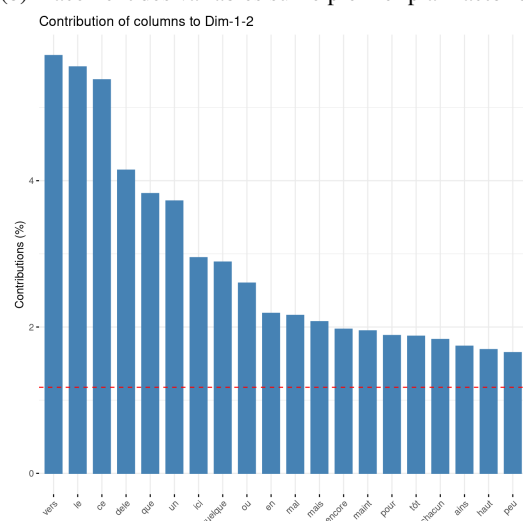
J-P Benzécri and M Meïmaris. 1996. Comparaison globale entre les œuvres de deux auteurs : Platon et Xénophon. *Cahiers de l'analyse des données*, 21(4) :403–430.



(a) Placement des textes sur le premier plan factoriel.



(b) Placement des variables sur le premier plan factoriel.



(c) Diagramme en barre vingt variables aux contributions les plus élevées pour ces deux axes.

FIGURE 1 – Résultats de l'analyse factorielle

- JP Benzecri. 1980. *L'analyse des données : tome 2, l'analyse des correspondances, éditions*. Bordas, Paris.
- Emmanuel Buron. 2006. Claude de Taillemont et les escriz de divers poètes à la louenge de Louïze Labé lionnoize. discussion critique de 'Louise Labé, une créature de papier', de Mireille Huchon. *L'information littéraire*, 58(2) :38–46.
- John Burrows. 2002. 'delta' : a measure of stylistic difference and a guide to likely authorship. *Literary and linguistic computing*, 17(3) :267–287.
- Florian Cafiero and Jean-Baptiste Camps. 2019. Why Molière most likely did write his plays. *Science Advances*, 5(11) :1–14.
- Florian Cafiero and Jean-Baptiste Camps. 2021. 'Psyché' as a Rosetta Stone? Assessing Collaborative Authorship in the French 17th Century Theatre. In *Proceedings of the Conference on Computational Humanities Research 2021*, volume 2989, pages 377–391. CEUR Workshop Proceedings.
- Florian Cafiero and Jean-Baptiste Camps. 2022. Who could be behind QAnon? Authorship attribution with supervised machine-learning.
- Florian Cafiero and Marie Puren. 2022. Claudine à l'atelier: l'impact de Willy et ses secrétaires sur l'écriture de colette. In *Humanistica*.
- Pramit Chaudhuri, Tathagata Dasgupta, Joseph P Dexter, and Krithika Iyer. 2019. A small set of stylometric features differentiates latin prose and verse. *Digital Scholarship in the Humanities*, 34(4) :716–729.
- Thibault Clérice, Vincent Jolivet, and Julien Pilla. 2022. Building infrastructure for annotating medieval, classical and pre-orthographic languages : the Pyrrha ecosystem. In *Digital Humanities 2022 (DH2022)*.
- Michele A Cortelazzo, George K Mikros, and Arjuna Tuzzi. 2018. Profiling Elena Ferrante: a look beyond novels. In *JADT 2018 : Proceedings of the 14th International Conference on Statistical Analysis of Textual Data*, pages 165–173.
- Guillaume Des Autels. 1550. *Repos de plus grand travail*. Guillaume Gazeau.
- Catherine Des Roches and Madeleine Des Roches. 1583. *Les secondes oeuvres de mesdames Des Roches*. N. Courtoys.
- Pernette Du Guillet. 1546. *Les Rithmes et poésies de gentile et vertueuse dame D. Pernette Du Guillet*. N. Courtoys.
- Maciej Eder. 2016. Rolling stylometry. *Digital Scholarship in the Humanities*, 31(3) :457–469.
- Maciej Eder. 2017. Short Samples in Authorship Attribution: A New Approach. In *DH*.
- Maciej Eder, Jan Rybicki, and Mike Kestemont. 2016. Stylometry with R: a package for computational text analysis. *The R Journal*, 8(1).
- Marc Fumaroli. 2006. Louise Labé, une géniale imposture. *Le Monde*.
- Simon Gabay. 2021. Beyond Idiolectometry? On Racine's Stylometric Signature. In *Proceedings of the Conference on Computational Humanities Research 2021*, volume 2989, pages 359–376. CEUR Workshop Proceedings.
- Simon Gabay, Thibault Clérice, Jean-Baptiste Camps, Jean-Baptiste Tanguy, and Matthias Gille-Levenson. 2020a. Deucalion, modèle français moderne (0.1.0).
- Simon Gabay, Thibault Clérice, Jean-Baptiste Camps, Jean-Baptiste Tanguy, and Matthias Gille-Levenson. 2020b. Standardizing linguistic data: method and tools for annotating (pre-orthographic) French. In *Proceedings of the 2nd International Conference on Digital Tools & Uses Congress, DTUC '20*, pages 1–7. Association for Computing Machinery.
- Graeme Hirst and Olga Feiguina. 2007. Bigrams of syntactic labels for authorship discrimination of short texts. *Literary and Linguistic Computing*, 22 :405–417.
- Mireille Huchon. 2005. *Louise Labé : une créature de papier*. Titre courant. Droz.
- Michel Jourde. 2017. Jacques Pelletier et Louise Labé: à propos de quelques hypothèses récentes. *Bibliothèque d'Humanisme et Renaissance*, 79(2) :299–311.
- Mike Kestemont. 2014. Function words in authorship attribution. from black magic to theory? In *Proceedings of the 3rd Workshop on Computational Linguistics for Literature (CLFL)*, pages 59–66.
- Mike Kestemont, Justin Stover, Moshe Koppel, Folger Karsdorp, and Walter Daelemans. 2016. Authenticating the writings of Julius Caesar. *Expert Systems with Applications*, 63 :86–96.
- Moshe Koppel and Yaron Winter. 2014. Determining if two documents are written by the same author. *Journal of the Association for Information Science and Technology*, 65(1) :178–187.
- Louise Labbé. 1555. *Euvres*. Jean de Tournes.
- Louise Labé. 2021. *Œuvres complètes*. Bibliothèque de la Pléiade. Gallimard. Ed. Huchon, Mireille.
- Sébastien Lê, Julie Josse, and François Husson. 2008. Factominer: an R package for multivariate analysis. *Journal of statistical software*, 25 :1–18.
- Olivier de Magny. 1553. *Les Amours*. Estienne Groulleau.
- Olivier de Magny. 1557. *Les Soupirs*. Vincent Sertenas.
- Olivier de Magny. 1559. *Les Odes*. André Wechel.
- Daniel Martin. 2006. Louise Labé est-elle «une créature de papier»? *Réforme, Humanisme, Renaissance*, 63(1) :7–37.
- Hermann Moisl. 2011. Finding the minimum document length for reliable clustering of multi-document natural language corpora. *Journal of Quantitative Linguistics*, 18(1) :23–52.
- Georgette de Montenay. 1571. *Emblèmes ou devises chrestiennes*. J. Marcorelle.
- Maurice Scève. 1544. *Délie, object de plus haute vertu*. Sulpice Sabon.
- Maurice Scève. 1547. *Saulsaye*. Jean de Tournes.
- Maurice Scève. 1562. *Microcosme*. Jean de Tournes.

Shachar Seidman. 2013. [Authorship verification using the impostors method](#). In *CLEF 2013 Evaluation labs and workshop—Working notes papers*, pages 23–26.

Claude de Taillemont. 1556. *La Tricarite*. Jean Temporal.

Pontus de Tyard. 1549. *Erreurs amoureuses*. Jean de Tournes.