

Genderqueer Marquis(e)

Computational Methods of Authorship Attribution for A French Tale

A Thesis presented to the faculty of the Université de Bourgogne Franche-Comté / Université de Franche-Comté, as part of the Sciences du Langage de l'Homme et de la Société department. In partial fulfillment of the requirements for the degree Master of Rare Books and Digital Humanities.

By Amanda Robin Hemmons, 2024

Abstract: My intent in this thesis project has been to apply computational methods of authorship attribution to the *Histoire de la Marquise-Marquis de Banneville*. This approach not only aims to clarify the authorship debate but also seeks to contribute to the broader discourse surrounding gender, collaboration, and the role of personal narrative in fairy tales and literary works of the late Seventeenth century. By grounding this analysis in quantitative evidence, I hope to offer a new perspective on the text, illuminating the possibilities of who may have authored the tale while underscoring the intricate history that has shaped the current understanding of it. Ultimately, this research endeavor will not only shed light on the complexities of the *Histoire* but also enrich the understanding of its place in the canon of fairy tales and its lasting significance in literary queer history.

Table of contents

Table of contents.....	2
Introduction.....	5
Textual Importance.....	8
◆ Folk tales, fairy tales, and contes galants.....	11
◆ The Histoire Itself.....	16
◇◇ Publication History.....	19
◇◇ Characters and Story.....	23
◆ The Potential Authors.....	24
◇◇ François-Timoléon de Choisy.....	25
◆◆◆ Collaboration between Perrault and Choisy.....	30
◇◇ Charles Perrault.....	32
◆◆◆ Collaboration between Perrault and L’Heritier.....	35
◇◇ Madame L’Heritier.....	37
◇◇ Other Configurations.....	40
◆ Whodunit? A brief history of authorship attribution.....	43
Computational Analysis.....	46
◆ What is NLP?.....	47
◆ Methodology.....	49
◇◇ Acquiring Text.....	51
◆◆◆ Multiple editions of the Histoire.....	52
◆◆◆ Charles Perrault texts for comparison.....	56
◆◆◆ Abbé de Choisy texts for comparison.....	58
◆◆◆ Madame L’Héritier texts for comparison.....	59
◇◇ Transcription: Transkribus and Tesseract.....	60
◇◇ Creating the Notebook.....	63
◇◇ Data Preprocessing.....	66
◇◇ Fast Stylometry for Burrough’s Delta.....	70
◆◆◆ Vocabulary Range.....	71
◆◆◆ Delta Results.....	73
◆ Findings.....	75
Conclusion.....	77
Digital Object.....	82
◆ Design & Layout.....	83
◆ Content.....	85
Bibliography.....	87

List of Figures/Tables

Figure 1.....	4
Figure 2.....	22
Figure 3.....	44
Figure 4.....	48
Figure 5.....	53
Figure 6.....	54
Figure 7.....	55
Figure 8.....	56
Figure 9.....	57
Figure 10.....	59
Figure 11.....	60
Figure 12.....	62
Figure 13.....	76
Figure 14.....	83
Figure 15.....	84
Figure 16.....	85
Figure 17.....	86
Table 1.....	63
Table 2.....	72
Table 3.....	73
Table 4.....	74

Genderqueer Marquis(e)



Figure 1. The abbé de Choisy dressed as a woman.
Illustration published in the magazine "The Museum of Families", in 1855.

Introduction

Queer stories and queer people are often described as a modern invention, as if they sprung fully formed from Madonna's bullet bra the same way Athena emerged from Zeus's head in mythology. In reality, we have stories dating back centuries about people with unusual gender presentation, and non-traditional romances that any modern reader would describe as queer.

Queerness and mythology, in fact, share deep roots, with ancient myths offering stories of fluid identities, transformations, and complex relationships that challenge the strict binaries often found in later societal norms. There are figures like Caenus¹, who transitioned from a female to a male form with Neptune's blessing, and many ancient pantheons that also included gods and goddesses with androgynous qualities. These examples show that fluidity in gender and identity was part of human storytelling long before modern labels existed. These themes of transformation, resistance, and self-discovery naturally carried over into the realm of fairy tales, where they have continued to evolve and resonate with audiences over time.

Mythology, folk tales, and fairy tales share deep connections in the storytelling tradition, even as they evolved to suit changing cultural needs. They have grappled with complex identities, often exploring questions of transformation and identity beyond a strict male-female binary, viewing them instead as manifestations of cosmic power. As societies grew and diversified, mythology adapted into more localized folk tales that shifted focus from gods and cosmic battles to human characters and everyday struggles. Folk tales retained elements of the mythic but brought them down to earth, focusing on peasants, animals, and mystical creatures that inhabited their listeners' worlds.

Yet even as the focus moved from gods to people, folk tales continued to include characters whose roles or identities did not conform to social norms. Figures like clever tricksters and gender-fluid creatures² embodied the same questioning of rigid identity boundaries as in myths. These stories often circulated within tight-knit communities, adapting to each culture's environment while retaining themes of queerness and unconventional self-presentation. In this sense, folk tales are mythology transformed: they keep myth's archetypal patterns and values while making them more accessible and relatable to everyday audiences. Folk tales upheld the

¹ Ovid, "Metamorphoses (Kline) 12, the Ovid Collection."

² di Laurea, "Loki's Gender Fluidity and Bisexuality in Norse Mythology and in Modern Adaptations."

myths' message of adaptability, presenting queerness not as an anomaly but as an intrinsic part of human nature³ and storytelling. By exploring identity through fantastic transformations or animal-human hybrids, folk tales demonstrated that diversity in self-expression was woven into the very fabric of communal storytelling.

With the advent of print culture, folk tales evolved further into literary fairy tales. Writers collected and adapted folk tales into a polished, literary genre that appealed to aristocratic tastes, while still preserving themes of otherness and transformation. Fairy tales began to incorporate structured plots, clear morals, and narrative closure; features that are rare in oral storytelling. While folk tales circulated orally and allowed for creative retellings, fairy tales became fixed in written form, with specific "authors" and authoritative versions. Yet despite their refined nature, fairy tales, like folk tales, continued to explore questions of personal identity, transformation, and acceptance—at times even more subversively, given their appeal to elite audiences.

Traditional fairy tales can seem to reinforce rigid gender roles and heterosexuality, but queerness and gender nonconformity can be found beneath the surface. Many classic stories feature characters who blur gender lines or defy societal expectations, often by taking on disguises or undergoing physical transformations. For example, the theme of the mermaid's desire to transcend her prescribed identity is most notably seen in Hans Christian Andersen's original 1837 version of *The Little Mermaid*⁴. In this version, the mermaid's yearning to become human reflects not only her romantic desire for the prince but also her deeper longing for a soul and a different existence, symbolizing a more complex transformation beyond mere physicality.

Many scholars have interpreted Andersen's tale as a metaphor for identity and self-discovery, with some readings suggesting a queer subtext. Andersen himself may have identified with the mermaid's sense of longing and desire to transcend boundaries, as some of his personal letters reveal unrequited love for men⁵. The mermaid's quest for an impossible transformation can be seen as a reflection of the struggle for acceptance and a place in a society that rigidly defines gender roles and expectations. In this way, her struggle to change her form and identity can be interpreted as resonating with themes of breaking free from societal limitations, including those related to gender and sexuality. Even the journey of self-discovery

³ Conner et al., *Cassell's Encyclopedia of Queer Myth, Symbol, and Spirit*, 22.

⁴ Andersen, "Hans Andersen's Fairy Tales. Second Series."

⁵ Fleming, "Hans Christian Andersen's Sexuality."

that many protagonists undergo echoes the journeys of those discovering and accepting their own queerness in a world that often tries to suppress it.

For marginalized people, especially those in the queer community, fairy tales offer an opportunity to reclaim narratives of transformation, otherness, and hope. Queer readers often relate to the fairy-tale trope of the outsider, finding kinship in the struggles of characters who long to live authentically in a society that doesn't fully accept them. Fairy tales, by their very nature, deal in fantasy, which provides marginalized groups with the space to reimagine worlds where they can be their true selves, where magic allows them to transcend the limitations of societal oppression.

Moreover, fairy tales often conclude with a sense of justice or a "happily ever after" that resonates deeply with those who have experienced exclusion. These stories hold out the promise of a better future—a narrative arc that parallels the longing for equality and belonging in real life. The archetypal simplicity of fairy tales make them easy to relate to, while their magical elements offer a hopeful escape from the harshness of reality. This is why queer people, alongside other marginalized communities, often find solace, empowerment, and a sense of possibility within these age-old stories.

Folk tales haven't disappeared, despite fairy tales' widespread influence. Oral storytelling is resilient, and queerness continues to evolve with them, emerging in urban legends, local myths, and even internet lore. Although literary fairy tales gave folk narratives a new life in print, they could never fully replace the oral tradition, which continues to transform, adapt, and reflect each generation's concerns. Together, mythology, folk tales, and fairy tales form a continuum that shows how stories evolve yet remain fundamentally similar—centuries later, they each echo the timeless themes of humanity's quest for meaning, identity, and belonging.

Textual Importance

I first encountered the *Histoire de la Marquise-Marquis de Banneville* in 2022 as a recent English translation in a collection titled the *Wonder Tales: six stories of enchantment*⁶. This

⁶ Marina Warner, *Wonder Tales*.

particular version of the *Histoire*, titled “The Counterfeit Marquise”, was brought to my attention via social media⁷, where it was described as a forgotten fairy tale with unexpectedly queer themes. At the time, I was in the midst of searching for a thesis topic for my Master’s program in Rare Books and Digital Humanities⁸ and this story immediately caught my attention. The idea of studying a fairy tale with such rich, queer themes seemed like the perfect opportunity to blend my personal and academic interests.

I was intrigued and I sought out the original French and discovered that—despite the attribution in a later published version—there is no *undisputed* author connected to the story. The timing of my discovery felt fortuitous. Here I was, studying in France, which allowed me to take advantage of access to French archives and rare collections. This proximity to key resources made the *Histoire* an ideal focus for my thesis, especially since I could conduct research directly on the original French text. The lack of attribution added a layer of mystery surrounding its creation and transmission. The absence of an identified author, coupled with the limited accessibility of these early texts, made it a prime candidate for computational analysis—a relatively new and underutilized approach in literary studies. I couldn’t help but think that modern methodologies, particularly those from the Digital Humanities, could offer new insights into the unresolved question of the *Histoire's* authorship.

So what started as a literary analysis of queer themes soon evolved into a more technical investigation. As I delved deeper, I found no existing evidence of computational analysis being used in the previous attribution attempts. There are other scholars who can write more eloquently than I on the subject of queer themes in fairy tales⁹, but given that Digital Humanities is still an emerging field, especially in literary studies, there are not many humanities scholars that have extensive programming experience, so I saw this as an opportunity to fill a gap.

By employing new techniques, I hope to contribute to the ongoing discourse surrounding the authorship of the *Histoire de la Marquise-Marquis de Banneville*. Although my analysis may not directly engage with the literary exploration of queer themes, I believe the act of highlighting a story like the *Histoire* will promote a broader understanding of queer stories within the academic community.

⁷ @laurasimonsdaughter, “The Counterfeit Marquise: A Literary Fairy Tale...”

⁸ “Master Rare Book and Digital Humanities.”

⁹ Seifert, “Introduction: Queer(ing) Fairy Tales.”

Using computational tools for author attribution became the core focus of my research. Digital Humanities, particularly in the context of authorship studies, offers methods such as stylometry—the statistical analysis of literary style—and which can provide clues about an anonymous text’s possible authorship by comparing it to other known works. As I researched the *Histoire* and its three potential authors, based on previous scholarly conjectures, my goal became to apply computational analysis to the *Histoire* and works attributed to these three authors, hoping to determine patterns that could point to a definitive author.

To accomplish this, I needed to gather and process texts attributed to these authors, as well as multiple versions of the *Histoire de la Marquise-Marquis de Banneville*. I then used techniques such as word frequency analysis, sentence structure comparison, and other computational methods to compare the *Histoire* against the works of its potential authors. This process involved building a dataset that allowed for a more detailed comparison, a task that required teaching myself more advanced coding techniques than those I had previously been familiar with. While my background in fine arts had provided me with some digital skills, this project pushed me to deepen my knowledge in areas such as Python programming, text mining, and data visualization.

The project also allowed me to draw upon existing resources in the Digital Humanities, such as The Programming Historian¹⁰, which offers tools for authorship attribution that became invaluable in my analysis. In the end, my thesis became a blend of computational analysis, rare book history, and French literary scholarship. By focusing on the authorship question, I aimed to shed light on the anonymous writer behind the *Histoire de la Marquise-Marquis de Banneville* while also demonstrating how modern digital tools can be applied to classic texts. Through this project, I hope to contribute both to the field of Digital Humanities and to the understanding of this uniquely queer, historically obscure fairy tale.

The *Histoire* also stood out to me because it existed in a unique historical and literary context. It was first published in the *Mercure Galant*, which was a pioneering periodical launched in 1672 and considered the peak of the *petite presse*¹¹ in Paris. The *Mercure Galant* played a crucial role in shaping public taste, particularly among elite readers, by showcasing sophisticated tales and current social trends. The *Mercure* catered particularly to educated

¹⁰ Laramée, “Introduction to Stylometry with Python.”

¹¹ Miller, “The French Periodical Press during the Reign of Louis XIV.”

women who were some of the most devoted readers of the period. For contemporary audiences, it also served as a window into cultural values and cultural ideas of the time, including evolving views on gender and class, which are subtly woven into the *Histoire de la Marquise-Marquis de Banneville*.

Seventeenth-century magazines like the *Mercure Galant* are now considered rare books due to their limited print runs and fragile nature, which have left few surviving copies. As early examples of serialized publications¹², these magazines were often read widely and then discarded, making original editions scarce and valuable to collectors and scholars today. The remaining issues, therefore, hold great value for scholars and collectors, as they offer a direct link to the literary and social trends of the past. Early periodicals like the *Mercure* represent one of the first forms of serialized publications, which set a precedent for modern journalism and literary magazines. For rare book scholars, these magazines offer insights not only into the content they held but also into the ways people consumed media during the time. The rarity of such early publications emphasizes the historical significance of the *Histoire de la Marquise-Marquis de Banneville*, enhancing its appeal as a subject for research and analysis.

In approaching my research on the *Histoire de la Marquise-Marquis de Banneville*, I recognized that studying such a text required an interdisciplinary perspective. With my academic background in Rare Books and Digital Humanities, I could explore both the physical rarity of early printed versions and the digital tools that allow us to analyze and preserve these works for future scholars. Combining traditional book history methods with digital analysis offered a fresh way to examine not only the text itself but also its context within the broader evolution of gender narratives in literature. Studying this *Histoire* in relation to other rare works from the same era allows me to delve into both the queer themes of the narrative and the broader historical forces that shaped its publication and reception.

Ultimately, researching the *Histoire de la Marquise-Marquis de Banneville* enabled a deeper understanding of how stories with progressive themes could exist and even flourish within the seemingly restrictive boundaries of 17th-century literary culture. This story's publication in the *Mercure Galant* reflects a nuanced acceptance of diverse narratives, suggesting that early modern audiences were more open to fluid portrayals of identity than we might assume. As I continue this research, I hope to contribute to the conversation surrounding both the

¹² Ibid.

history of queer themes in literature and the value of rare publications in shaping our understanding of the past.

◆ Folk tales, fairy tales, and *contes galants*

While folk tales are a world-wide concept that predate written language, the modern concept of a fairy tale originates in Western Europe. Codified over time, they are not long novels, but rather short narratives with a certain expected structure¹³, inspired by oral and popular tradition. The opening line, “once upon a time,” is meant to put the reader in the mindset of a different time and place. There is a clear hero or heroine to the story who must overcome a series of trials; they have helpers, and these helpers are often fairies or animals. And, while the term “fairy tale” implies the presence of magical creatures, fairies aren’t actually a requirement of the genre, which is better characterized by acts of wild imagination and strong symbolism¹⁴.

The modern concept of a fairy tale as we know it today began to take shape during the Renaissance, a period of intellectual and cultural revival in Europe. This era saw a renewed interest in literature, art, and the exchange of ideas, leading to the creation and spread of many of the most enduring and beloved fairy tales. Among the key figures who contributed to this rich tradition were Italian authors Giovanni Francesco Straparola and Giambattista Basile. Their works in the Sixteenth and Seventeenth centuries laid the foundation for stories that continue to captivate audiences.

Straparola, often considered the father of the literary fairy tale, wrote the earliest known version of *Puss in Boots* in his *Le piacevoli notti* (The Pleasant Nights) in 1550¹⁵. This tale, titled *Costantino Fortunato*, introduced readers to the clever, anthropomorphic cat who helps his master achieve wealth and status through cunning and deception. Straparola’s collection was notable not only for its fairy tales but also for its framing story structure, in which different narrators recounted stories over several nights—an innovation that influenced later European storytellers.

¹³ Besson, “From the Aristocracy to the Playground.”

¹⁴ Warner, *Once upon a Time*.

¹⁵ Bottigheimer, *Fairy Tales*, 12.

This narrative framework was likely inspired by Giovanni Boccaccio's *The Decameron*¹⁶, written nearly two centuries earlier in 1353, a work that pioneered the use of a framing device to unify a set of disparate stories within a single overarching narrative. In *The Decameron*, a group of ten young people flees to the countryside to escape the Black Plague, and each day, they entertain each other by telling tales of romance, wit, and deception.

Boccaccio's work laid the groundwork for future storytelling traditions by illustrating how an anthology of tales could be woven together with a unifying theme or scenario, turning a collection into something greater than its individual parts. This structure established a precedent for later authors like Straparola, who adapted the idea of framing tales to create a cohesive, multi-layered narrative.

Straparola's narrators, like Boccaccio's, exchange stories over several nights, establishing a tradition in which fictional storytellers are as central to the experience as the stories themselves. Following Straparola's example, authors like Giambattista Basile and, later, Charles Perrault, continued to use this framework in their collections, including *The Tale of Tales* and the *Histoires ou contes du temps passé*. Each added layers of cultural context and social themes, such as transformation, cunning, and upward mobility, that resonated with audiences of their respective times.

Basile's work, in particular, introduced darker, folkloric elements that foreshadowed the later themes of the Brothers Grimm, who drew from these traditions to compile their own collection of tales. Giambattista Basile, writing in the early Seventeenth-century, followed in Straparola's footsteps by compiling and refining a collection of stories steeped in local folklore. His *Lo cunto de li cunti*¹⁷ (*The Tale of Tales*), published posthumously between 1634 and 1636, included some of the earliest known European versions of what are now iconic fairy tales. Among his works were *Cagliuso*, his own take on *Puss in Boots*, and *Petrosinella*, which would later evolve into the story of *Rapunzel*. Basile also wrote *La gatta Cenerentola*, an early version of *Cinderella* in which the heroine is aided not by a fairy godmother but by a magical, protective tree. His stories, rich in local dialect and dark humor, were intended primarily for adult readers, reflecting the fairy tale's origins as a genre for adults rather than children.

¹⁶ Boccaccio, *The Decameron*.

¹⁷ "SurLaLune Fairy Tales."

The popularity of these Italian stories grew as they reached other parts of Europe, particularly in France, where they gained further refinement and notoriety in the literary salons of the late Seventeenth-century. These salons, often hosted by aristocratic women¹⁸, were influential cultural hubs where writers, intellectuals, and nobility gathered to discuss literature, philosophy, and social issues. Within this context, fairy tales evolved from simple folk narratives into sophisticated stories that served as both entertainment and subtle social commentary. The salons provided a rare space for women to contribute to literature and intellectual discourse, allowing them a platform to explore and challenge ideas about gender, power, and class.

Marie-Catherine d'Aulnoy, one of the most renowned writers from these salons, is often credited with coining the term "fairy tale" ("conte des fées")¹⁹ in her works, which often featured strong female characters and fantastical elements. D'Aulnoy's work also introduced iconic elements to the genre, such as the figure²⁰ of Prince Charming, which would later become emblematic of fairy tale romance and adventure. Her tales, along with those of other salonnières like Charlotte-Rose de Caumont de La Force and Henriette-Julie de Murat²¹, were a transformative force, shifting the fairy tale's focus from simple moral lessons to complex narratives that included themes of love, loyalty, and personal growth. These writers expanded the fairy tale's thematic scope, subtly embedding reflections on the constraints and possibilities for women in French society.

D'Aulnoy's contemporary, Charles Perrault, also made significant contributions to the genre. In fact, he is one of the most famous figures to emerge from this period. His collection *Les Contes de ma Mère l'Oye*²² (Tales of Mother Goose), published in 1697, remains a cornerstone of the fairy tale tradition. Stories such as *Cendrillon* (Cinderella), *Le Petit Chaperon Rouge* (Little Red Riding Hood), and *La Belle au bois dormant* (Sleeping Beauty) gained widespread popularity through Perrault's versions, which were notable for their moral lessons and elegant prose.

The fairy tale tradition eventually spread beyond France and Italy, influencing writers across Europe. Including Germany, where the Brothers Grimm played a pivotal role in preserving oral folktales and bringing them to a wider audience. Their collection, *Kinder- und*

¹⁸ Besson, "From the Aristocracy to the Playground."

¹⁹ Aulnoy, *Contes Des Fées*, 1.

²⁰ Aulnoy, *Contes Des Fées*, 181.

²¹ Zipes, *The Great Fairy Tale Tradition*, 830.

²² Perrault, "Contes ou Histoires du temps passé - Les Contes de ma Mère l'Oye."

*Hausmärchen*²³ (Children’s and Household Tales), published in the early 19th century, gathered stories from local German traditions, many of which had been passed down orally for generations. Their versions of *Aschenputtel* (Cinderella), *Rapunzel*, and *Dornröschen* (Sleeping Beauty) became part of the cultural fabric of Western storytelling, with their darker, more morally complex narratives standing in contrast to the refined French versions.

The popularity of these stories led them to be translated and widely read in English, helping to establish the genre in English literary culture. As literacy rates increased and books became more accessible, fairy tales began to appeal to a growing middle-class audience, particularly children, and writers²⁴ blended fairy tales with moral lessons, contributing to the genre’s early prominence in English children’s literature. The Grimm Brothers’s collection, finally translated into English²⁵, also captivated readers.

Meanwhile English authors like George MacDonald created original works that incorporated traditional fairy tale elements. His novel, *The Princess and the Goblin*,²⁶ is an example how English writers crafted stories using the established themes of spiritual and mystical transformation.

In the Nineteenth-century, the genre gained further popularity, particularly with the influence of the Romantic movement, which celebrated²⁷ folklore, imagination, and the supernatural. Inspired by this we see another significant contribution to English fairy tale culture during the late 19th century, Andrew Lang’s *Fairy* books. Beginning with *The Blue Fairy Book*²⁸ in 1889, Lang compiled and adapted fairy tales from across Europe and beyond, creating a multi-volume series that introduced countless readers to stories from various cultures, including those familiar tales earlier imagined in Italy, *Puss in Boots* and *Cinderella*. Lang’s *Fairy* books became an essential part of Victorian children’s literature, offering beautifully illustrated and engaging stories that helped solidify the fairy tale’s enduring appeal in England.

Through these various iterations and adaptations, fairy tales evolved from their Renaissance roots into the almost-universal stories we recognize today. They transcended borders, appealing to both adult readers and children, as well as to a wide variety of

²³ Grimm and Grimm, *Kinder- und Haus-Märchen*.

²⁴ Carpenter and Prichard, *The Oxford Companion to Children’s Literature*, 73-74.

²⁵ “Grimm’s Fairy Tales | History & Best-Known Works | Britannica.”

²⁶ MacDonald, *The Princess and the Goblin. With Numerous Illustrations*.

²⁷ Taylor, “The Evolution of the Romantic Fairy.”

²⁸ Lang, Ford, and Hood, *The Blue Fairy Book*.

backgrounds. They became a vehicle for both entertainment and the exploration of deeper themes. These stories continue to resonate with audiences still today, illustrating the enduring power of fairy tales in shaping our imagination and culture.

In the Anglophile literary tradition today, fairy tales are a broad category and include a variety of story types, including those that might be classified differently in other cultures. For example, in French literature, the “conte galant” (or “gallant story”) refers specifically to tales set in aristocratic or royal court settings, often focusing on themes of love²⁹, courtship, and the social intrigues of the aristocratic class, rather than magical or fantastical elements. Unlike traditional fairy tales with overtly magical elements, contes galants are usually more grounded in real-world settings, where the focus lies on romantic entanglements and the power dynamics of elite society. However, they share with fairy tales a sense of heightened reality, often depicting idealized characters and scenarios that carry a dreamlike quality.

Despite the lack of overt magic, contes galants align with the fairy tale tradition by exploring universal human emotions through imaginative storytelling techniques. These tales frequently involve elements such as disguises, mistaken identities, and unexpected plot twists—devices that, while grounded in human cunning rather than supernatural intervention, create a sense of intrigue and suspense reminiscent of fantastical tales. Through these devices, contes galants weave a fine line between reality and fantasy, engaging readers with narratives that feel familiar yet artfully detached from everyday experience.

Additionally, contes galants, like fairy tales, often serve as allegories for moral or social lessons, reflecting and subtly critiquing the society of their time. Characters in these stories navigate both the pleasures and constraints of courtly life, where rules of decorum, honor, and power interplay with personal desire. The aristocratic backdrop provides an ideal setting to explore themes of loyalty, ambition, and social status—core issues that, in their complexity, offer a mirror to human nature. Thus, while contes galants may seem more restrained in magical content, their blend of social realism and narrative flair allows them to sit comfortably within the broader fairy tale tradition, connecting with readers' imaginations while also commenting on human relationships and societal norms.

²⁹ Weatherley, “Avatars of Gendered Societal Constructs in Seventeenth-Century Contes de Fées,” 17.

◆ The Histoire Itself

The story, *Histoire de la Marquise-Marquis de Banneville*, first appeared in 1695 in the *Mercure galant*, a prominent French literary magazine that set the stage for popular storytelling and social commentary in late seventeenth-century France. The *Mercure galant*³⁰ was the brainchild of Donneau de Visé, a historian, author, playwright, and literary critic. During the late Seventeenth-century, Paris had a thriving publishing scene with over a thousand gazettes in circulation, but the *Mercure galant* stood out for its wide appeal and influence. Nearly every significant author of the era was published in its pages, from celebrated figures like Charles Perrault, known for his fairy tales, to poet Jean de La Fontaine. From its inception in 1672 until 1714, the *Mercure galant* featured a diverse array of content, including poetry, short stories, sermons, and even fashion critiques, making it a cultural touchstone in literary and social circles.

When Donneau de Visé launched the *Mercure galant* in 1672, he aimed to blend entertainment with intellectual discourse, catering to an emerging audience of educated readers who desired news and cultural insights alongside the amusements of poetry, stories, and gossip. Visé's creation offered a variety of content that appealed to a broad spectrum of readers, ranging from elite literary circles to the educated bourgeoisie. The magazine included poetry, short stories, religious sermons, and social critiques, making it a go-to source for understanding the latest in French culture. Visé even anticipated modern publishing trends by incorporating fashion commentary, turning the *Mercure galant* into one of the earliest lifestyle magazines and a cultural touchstone for Parisians. The success of the *Mercure galant* was due not only to Visé's editorial talent but also to his collaboration with notable figures.

In 1680, playwright and poet Thomas Corneille, the brother of the renowned dramatist Pierre Corneille, joined Visé in shaping the magazine's editorial voice. Corneille's influence added intellectual depth and literary sophistication to the publication, further enhancing its appeal. Their combined editorial oversight helped position the magazine as a tastemaker in Parisian society, a platform where literary trends and new ideas could be tried and tested. This partnership marked a high point in the magazine's history, as Corneille's dramatic experience and Visé's knack for popular content struck a balance that resonated well with readers.

³⁰ Moureau, "LE MERCURE GALANT | Presse18."

Adding another layer to the *Mercure galant's* legacy, Donneau de Visé's nephew was none other than Bernard Le Bovier de Fontenelle, who would become an influential philosopher and essayist. Fontenelle contributed regularly to the *Mercure galant*, bridging the world of literature and science, and helping to shape public opinion on emerging scientific ideas. Fontenelle's connection to Visé and his role as a contributor further enriched the magazine's reputation, bringing intellectual rigor to its content while appealing to readers' curiosity about new discoveries. This unique blend of literature, science, and culture made the *Mercure galant* a true forerunner of interdisciplinary media.

In 1714, after Donneau de Visé's death, the magazine was renamed the *Mercure de France*³¹ and underwent several editorial changes, occasionally pausing publication for years at a time. Although it never regained the prominence it held during Visé's and Corneille's tenure, the *Mercure de France* remained a respected publication and an important part of France's literary heritage. Various editors and owners took the helm over the years, causing shifts in the magazine's focus and leading to occasional interruptions in publication, and ceasing at last in 1965 to be absorbed into the Éditions Gallimard publishing group. But its impact on shaping public opinion and culture remained evident long after Visé's death, as it influenced perspectives on everything from politics to literature.

It was within this influential magazine that the first version³² of the *Histoire de la Marquise-Marquis de Banneville* was published in February 1695. The initial publication generated considerable interest, prompting a reprint the following year. This reprint, presented in two parts, included substantial additions that nearly doubled the story's original length. The expansion of the tale speaks to its popularity; by revisiting and enlarging it, the editors showed that the story had captured the public's imagination, presenting it as a tale worth preserving and elaborating upon. This suggests that the tale captured readers' imaginations, contributing to its continued interest as both a literary and historical artifact.

This growing interest culminated in its third publication by d'Houry in Paris. Unlike the previous versions published in the *Mercure galant*, this edition was released as a standalone text and included two additional scenes featuring the character of the Marquis de Bercourt. Notably,

³¹ Moureau, "LE MERCURE GALANT | Presse18."

³² "Histoire de la Marquise Marquis de Banneville."

this publication did not contain an editor's note beyond the Privilege du Roy³³ and still lacked any indication of a possible author, adding to the ongoing confusion regarding authorship.

The story's continued appeal as both a literary and historical artifact underscores the *Mercure galant's* role as a cultural hub. Stories like *Histoire de la Marquise-Marquis de Banneville* illustrate the magazine's ability to shape public taste, provide entertainment, and stimulate intellectual curiosity among readers. By merging literary innovation with societal reflections, the *Mercure galant* and its contributors laid the groundwork for future literary magazines and helped establish the periodical as a critical medium for storytelling in French literary tradition.

³³ Official endorsement from the crown, standard for books legally published at the time.

❖ ❖ Publication History

1695

Mercure galant pp. 14-101, Paris³⁴

First edition. Printed in the February edition of the magazine. No name attached, but prefaced with an author's note implying it to be written by a young woman to be read by other young women. (7,085 words)

1696

Mercure galant pp. 171-238 and 85-185, Paris³⁵

Second edition. Part one printed in August, with an editor's note mentioning the woman author; part two printed in September, with an added in-text discussion/advertisement of *la Belle au bois dormant* by Charles Perrault. Additional background story of Prince Sionad added. (13,374 words)

1723

d'Houry, Paris³⁶

Third edition. Published as a standalone text with two additional scenes featuring the Marquis de Bercourt. No author's note, no editor's note beyond the Privilege du Roy. (13,868 words)

1928

Mercure de France, Paris³⁷

Reprint of the first edition. Prefaced by Madame Jeanne Roche-Mazon's scholarly paper where she outlined her belief that the *Histoire* was a collaboration between Charles Perrault and the Abbé de Choisy.

³⁴ "Histoire de la Marquise Marquis de Banneville," *Mercure galant*, February 1695.

³⁵ "Histoire de la Marquise Marquis de Banneville," *Mercure galant*, August 1696; "Histoire de la Marquise Marquis de Banneville," *Mercure galant*, September 1696.

³⁶ *Histoire de la marquise, marquis de Banneville*, 1723.

³⁷ "Histoire de la Marquise Marquis de Banneville," *Mercure de France*, February 1928.

1928

La Centaine, Paris³⁸

A standalone text and a reprint of the first edition. Includes an editor's note presenting Madame Roche-Mazon's conclusion of the Perrault-Choisy collaboration.

1973

Peter Owen, London³⁹

Published as *The Transvestite Memoirs of the Abbé de Choisy*, this next version of the *Histoire* was printed along with the abbé's memoirs. This was the first official English translation (of the first edition) by R.H.F. Scott, and included an editor's note that he believed it to be more likely Choisy's work than Perrault's.

1994

Peter Owen, London and Chester Spring, PA

A reprint of *The Transvestite Memoirs* using Scott's translation of the first edition and published for a wider market, in the United States as well as London.

1994

Chatto & Windus, London

Six french fairy tales collected by Marina Warner into a book called the *Wonder Tales*⁴⁰; one of which was an English translation of the *Histoire*'s first edition, titled "The Counterfeit Marquise" and translated by Ranjit Bolt. Warner's introduction repeats the idea that the story is a collaboration between Charles Perrault and the abbé de Choisy.

³⁸ Perrault, *Histoire de la Marquise-Marquis de Banneville par Charles Perrault et l'abbé de Choisy*.

³⁹ De Choisy, *The Transvestite Memoirs and the Story of the Marquise-Marquis De Banneville*.

⁴⁰ Marina Warner, *Wonder Tales*.

1995

Editions Ombres, Toulouse⁴¹

Mémoires de l'abbé de Choisy habillé en femme : suivi de, Histoire de la marquise-marquis de Banneville. Like *The Transvestite Memoirs*, this is a copy of Choisy's memoirs alongside the *Histoire*. This edition seems to assume Choisy alone wrote the *Histoire*.

1996

Vintage, London and Farrar, Straus and Giroux, New York

A *Wonder Tales* reprint of "The Counterfeit Marquise" expanded to include the USA market.

2002

Random House & Overdrive

The first ebook edition of Marina Warner's *Wonder Tales* collection.

2004

Oxford University Press, Oxford

Another *Wonder Tales* reprint, based on the American edition.

2004

Modern Language Association of America, New York⁴²

They published both the French and English at the same time, translated by Steven Rendall and Joan E. DeJean. A reprint of the second edition, with the two added scenes from the third edition included as appendices. These versions of the *Histoire* include an introduction by DeJean discussing the possibilities of authorship being Charles Perrault, Madame L'Héritier, the abbé de Choisy, or some combination thereof.

⁴¹ Choisy, *Mémoires de l'abbé de Choisy habillé en femme ; suivi de Histoire de la marquise-marquis de Banneville*.

⁴² Perrault, Choisy, and L'Héritier de Villandon, *The story of the Marquise-Marquis de Banneville*.

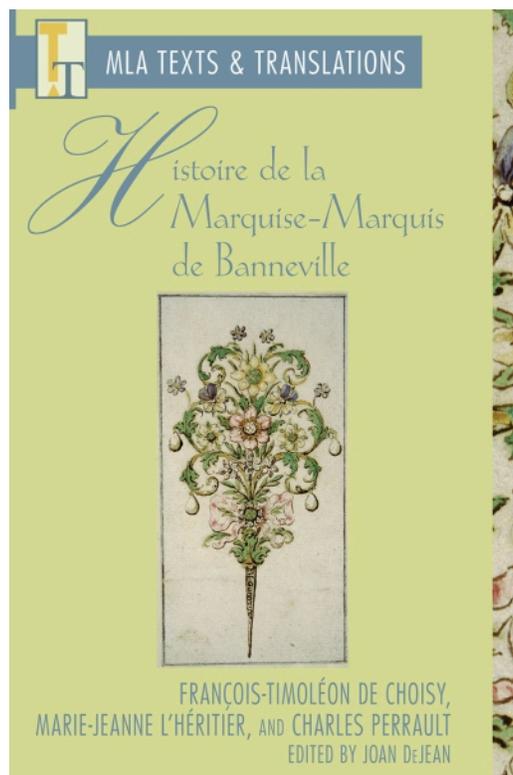


Figure 2. Book cover of the MLA edition of the *Histoire de la Marquise-Marquis de Banneville*.

❖ ❖ Characters and Story

The Characters

Mariane

The main character of our story, often called la petite Marquise. An AMAB⁴³ character raised unknowingly as a girl.

Marquise de Banneville

Mariane's mother.

Comtesse d'Aletref/de Variere

An old friend of the Marquise and takes the role of Mariane's "fairy" godmother. The name was changed from Aletref to Variere in the third edition.

Mme d'Aletref/de Variere

The Comtesse's daughter and Mariane's friend.

Prince Sionad

A young foreign prince known to be both a great warrior and a crossdresser.

Marquise de Bercour(t)

A soldier and Mariane's love interest, eventually revealed to have been AFAB⁴⁴. Bercour in the first edition and Bercourt thereafter.

Counselor

The Marquise de Banneville's brother and Mariane's eventual guardian.

The Story

Determined to keep her son from dying in war the way her husband did, the Marquise de Banneville raises her child as a girl named Mariane.

When Mariane is a teenager, she and her mother go to Paris where she is taken under the wing of the Comtesse. Mariane encounters Prince Sionad dressed as a woman at a ball and is shocked to discover he is a crossdresser. This is her first encounter with non-standard gender presentation.

Mariane also meets the Marquis de Bercourt and they quickly strike up a romance, although the Marquis is unusually resistant to the idea of marriage.

When Mariane tries to press the issue, her mother confesses the truth of Mariane's birth. Mariane is heartbroken over this, but eventually determines that she wants to marry the Marquis anyway.

Her mother, the Marquise, has died of illness in the meanwhile and her uncle consents to the union, motivated by the expectation of a childless marriage that would preserve his heirs' inheritance.

On their wedding day, an unexpected twist unfolds when the Marquis discloses that she was born a woman. In turn, Mariane reveals her own truth of having been born a boy. The couple decides they will continue to embrace their identities and continue living as the gender they have always known. Their union ultimately leads to the birth of a child, defying her uncle's expectations, and they lived happily ever after.

⁴³ Assigned male at birth.

⁴⁴ Assigned female at birth.

◆ The Potential Authors

The *Histoire de la Marquise-Marquis de Banneville* has been attributed to three different authors over the three centuries since its inception, although few definitive opinions have been preserved in accessible formats. The majority of scholarly discourse surrounding the authorship has emerged in the last century.

In 1906, Paul Bonnefon published *Les dernières années de Charles Perrault*⁴⁵ (The Last Years of Charles Perrault), suggesting that the *Histoire* was written by a female author rather than the widely assumed Charles Perrault himself. Most likely Madame L'Héritier, who was not only a relative of Perrault but also an emerging literary figure in her own right.

Later, in 1928, Madame Jeanne Roche-Mazon contributed to the conversation by proposing that the *Histoire* was in fact a collaborative effort⁴⁶ between Charles Perrault and the abbé de Choisy, another literary figure of the period. Choisy had a unique background that could have brought a stylistic and thematic layer very different to Perrault to the work. Roche-Mazon's assertion introduced a new layer to the authorship debate, suggesting that the text may be the result of combined literary talents rather than the sole creation of one author. Despite these discussions, a definitive consensus regarding the authorship of the *Histoire* remained elusive.

Somewhat more recently, the Modern Language Association of America published the *Histoire de la Marquise-Marquis de Banneville* as a standalone text in 2004, alongside *The Story of the Marquise-Marquis de Banneville*⁴⁷, translated by Steven Rendall. This edition includes an introduction by Joan DeJean, which examines the potential authorship of the *Histoire*, considering Charles Perrault, Madame L'Héritier, the abbé de Choisy, as well as the potential for some combination thereof.

In the next section I will describe the literary evidence presented by these scholars. While their arguments provide substantial insight, they primarily rely on historical context and close readings of the text itself. Importantly, no statistical or computational analyses have been conducted to further investigate these claims. This gap in research highlights an opportunity for a computational approach to authorship studies, particularly in the context of the *Histoire*. By

⁴⁵ Bonnefon, "Les Dernières Années de Charles Perrault," 623.

⁴⁶ Roche-Mazon, "An unexpected collaboration in the 17th century: the abbé de Choisy and Charles Perrault," 528.

⁴⁷ Perrault, Choisy, and L'Héritier de Villandon, *The story of the Marquise-Marquis de Banneville*.

employing modern techniques of text analysis, I hope to unearth patterns and characteristics that may suggest specific authorship or reveal collaborative influences among the potential authors.

❖ ❖ François-Timoléon de Choisy

François-Timoléon de Choisy⁴⁸ was raised on intimate terms with the French court, which shaped much of his life and career. Born in Paris in 1644, his father was part of the duke of Orléans' household and his mother close friends with Anne of Austria, mother of King Louis XVI. Encouraged by his mother, Choisy chose a religious path over a military career, joining the church and studying philosophy and theology at the Sorbonne. In 1663, he secured the role of abbot of Saint-Seine in Burgundy, a position that provided him with both spiritual responsibility and a stable income.

Despite his mother's urging to pursue a disciplined life, Choisy soon found himself drawn to a life of luxury and adventure. After his mother's death in 1669, Choisy indulged in the high society and decadence of the time, eventually gambling away his money in Venice. Forced to return to France to rely on his ecclesiastical income, he attempted to find a renewed purpose, spending time in Rome with his friend Cardinal de Bouillon and forming connections with influential figures like the Bishop of Valence. It was only after a severe illness that Choisy reevaluated his life, leading him to shift focus towards religious and scholarly pursuits. He accepted a diplomatic assignment to Siam,⁴⁹ joining an embassy sent by Louis XIV. This journey resulted in a travel memoir detailing his experiences in Siam and marked a shift toward his role as a respected historian.

Upon his return to France, Choisy continued to write, contributing numerous religious and historical works, including an eleven-volume *Histoire de l'Église*⁵⁰ (History of the Church). His writing skills and connections earned him the position of literary director for Madame de Lambert, a prominent figure in Parisian literary circles. His reputation as a scholar grew, and he joined the prestigious Académie française. There, he collaborated with Charles Perrault on

⁴⁸ Choisy, *Mémoires de l'abbé de Choisy habillé en femme*, 2.

⁴⁹ Metz Howard, "Professor Casts Doubt on One of History's Greatest Cross-Dressing Memoirs."

⁵⁰ Choisy, *Histoire de l'Église. Tome 1 / , Par M. l'abbé de Choisy...*

*Opuscules sur la langue française*⁵¹ and contributed to the famous dictionary, a project reflecting his enduring interest in refining and promoting French literature and language.

So on the one hand, Choisy was a gifted scholar, and on the other, he is known not only for his intellectual contributions but also for his unconventional personal life, specifically his well-documented cross-dressing. His personal accounts, *Mémoires pour servir à l'histoire de Louis XIV* (Memoirs to serve the History of Louis XIV) and *Histoires de Madame la comtesse des Barres* (The Story of the Countess of Barres), edited and later published as *Mémoires de l'abbé de Choisy habillé en femme* (Memoirs of the Abbot de Choisy Dressed as a Woman), detail his experiences dressing as a woman, a habit he began in childhood. His mother was the one who first dressed him in women's clothes, a practice she encouraged until he was eighteen. After her death, Choisy continued wearing her jewels and fine clothing, blending his identity with a feminine guise that he presented in both public and private life. This upbringing draws an interesting parallel with the character of Mariane de Banneville in the *Histoire de la Marquise-Marquis de Banneville*, whose mother similarly tries to shield her from the hardships of war by directing her toward a safer, more restricted role in society.

Some scholars believe that the similarities between Choisy's life and the fictional character Mariane de Banneville hint at an autobiographical element within the *Histoire*, potentially written by Choisy himself or someone inspired by him. In fact, Choisy's biographer, de Perceflour, even speculated in a 1920 edition of Choisy's memoirs that there might be a missing manuscript describing the Marquise-Marquis de Banneville. De Perceflour suggested⁵² that this missing section could contain Choisy's own ideas and contributions to the character, introducing her as an episodic figure whose storyline closely aligns with Choisy's real-life experiences.

Il y a certainement entre ce chapitre et le précédent une lacune causée par la destruction d'un des fragments du manuscrit original. C'est l'introduction d'un personnage épisodique, la marquise-marquis de Banneville dont le caractère a dû se développer dans des pages qu'on n'a pas retrouvées.

(There is certainly a gap between this chapter and the previous one caused by the destruction of one of the fragments of the original manuscript. It is the introduction of an episodic character, the Marquise-Marquis de Banneville, whose character must have developed in pages that have not been found.)

⁵¹ Roche-Mazon, "An unexpected collaboration in the 17th century: the abbé de Choisy and Charles Perrault," 521.

⁵² Choisy, *Mémoires de l'abbé de Choisy habillé en femme*, 147.

De Perceflour's theory that the character might have been based on Choisy's unique life experience gives weight to the hypothesis of Choisy's involvement in writing the tale.

That being said, one can't base our assumptions on a piece of manuscript that has never been seen. Although there are striking parallels between Choisy's life and the events in the *Histoire*, this does not necessarily make the story autobiographical. It is equally possible that someone else created the character of Mariane based on rumors of Choisy's public life. Choisy was notorious within French aristocratic circles, and his penchant for dressing as a woman, even in religious settings, may have provided inspiration for the Marquise-Marquis character. This notion raises questions about how widely known Choisy's lifestyle actually was during his lifetime and how it might have influenced those around him.

Choisy's own memoirs suggest he was open about his cross-dressing, describing occasions when he would dress as a widow to attend church or be seen at social gatherings. According to Choisy's accounts, he continued this practice throughout his life, even into his eighties. While Choisy's detailed accounts have been largely accepted as factual, some scholars have questioned the authenticity of his stories.

Paul Scott, an associate professor of French, has pointed out that certain aspects⁵³ of Choisy's memoirs appear anachronistic. For instance, Scott notes that Choisy describes wearing a *stinquerque*⁵⁴, a particular type of garment that did not become popular until 1692, long after some of the exploits described. This raises the possibility that some of Choisy's memoirs may have been embellished or retrospectively constructed.

The combination of Choisy's public cross-dressing, his mother's role in shaping his identity, and his life within the highest circles of French society provides fertile ground for speculation regarding his connection to the *Histoire*. The presence of a gender-fluid character like Mariane de Banneville in a French literary work of that time period is itself intriguing, and it is possible that Choisy, either directly or indirectly, influenced the character. The *Histoire de la Marquise-Marquis de Banneville* might represent a subtle acknowledgment of Choisy's life by an unknown author or serve as an early example of gender exploration in French literature, consciously or unconsciously inspired by him. Several details in the *Histoire de la*

⁵³ Scott, "Authenticity and Textual Transvestism in the Memoirs of the Abbe de Choisy."

⁵⁴ A type of fabric band, originating after the Battle of Steenkerque in 1692, used to cover the skin exposed by low necklines and considered a precursor to modern scarves and neckties.

Marquise-Marquis de Banneville hint at Choisy's authorship through recognizable parallels to his life and social circle.

For instance, the character of the Countess d'Alettef in the story is likely a playful reference to Madame de La Fayette, a close friend of Choisy's. In both the *Histoire* and real life, these women gave the same ironic advice to effeminate young men: if they were going to wear earrings, they might as well embrace femininity fully and dress as women. In the *Histoire*, Countess d'Alettef gives this advice to the Marquise about the Marquis de Bercourt, while Madame de La Fayette offered⁵⁵ the same to Choisy himself. Additionally, Prince Sionad, a character in the *Histoire*, is widely interpreted as an anagram of "Adonis," the mythical figure renowned for his beauty, likely representing the Duke of Orléans, known for occasionally attending events in women's attire. The Duke was part of Choisy's social world and present at court, reinforcing this connection. Furthermore, Mariane's uncle in the *Histoire* shares striking similarities with Choisy's own uncle, who also served as a counselor. These interwoven details, drawn from Choisy's life and relationships, strongly support the case for his authorship of the *Histoire*.

In his memoirs, Choisy references the *Histoire de la Marquise-Marquis de Banneville* directly, though this mention only complicates the question of authorship further. He recounts an instance⁵⁶ in which, in his guise as Mme de Sancy, he entertained a parish priest and Mademoiselle Dupuis at his home. They read the latest *Mercure galant*, where a recent story had caught Choisy's eye, though he describes Mariane as a man who wanted to live as a woman, adorning himself in elegant dresses, jewelry, and hairpieces, much like Choisy himself did. "I see well," Choisy reportedly said, "that it resembles me, but I do not know if I should be angry about it." Dupuis replied, encouraging him not to take offense, suggesting that the story was flattering in its own way. She even joked that she wished the author had named Choisy outright so that his notoriety might spread further.

However, this account presents a chronological dilemma. Choisy's experiences as Mme de Sancy predate the publication of the *Histoire* by nearly three decades, raising doubts about the veracity of his memory or the memoir itself. Scholar Paul Scott points to this as a glaring anachronism, noting that if Choisy embellished or perhaps fictionalized this encounter, it was not

⁵⁵ Choisy, "Aventures de l'abbé de Choisy habillé en femme," 75.

⁵⁶ *Ibid*, 19.

uncommon in memoirs of the time, particularly for someone with Choisy's literary and performative flair. Scott suspects⁵⁷ that Choisy's recollections may have been editorialized to enhance his persona, a blend of eccentricity and self-reflection. Yet, Scott still supports Choisy as one of the potential authors of the *Histoire*, if only due to his open reference to it.

Scott is somewhat dismissive of Choisy's claims to have continued dressing as a woman into his later years, doubting that he would do so once he was no longer able to convincingly "pass" as young and beautiful. However, this skepticism seems to overlook the possibility that Choisy's cross-dressing was motivated less by external appearance or social acceptance and more by personal enjoyment. While Choisy may have embellished parts of his memoirs for effect, such embellishments don't necessarily discount the underlying truth of his experiences. Rather, Choisy's cross-dressing seems to stem from a genuine affinity for presenting himself as a woman, whether or not he maintained the illusion of youth or beauty. This nuanced perspective, often overshadowed by conventional interpretations, suggests that Choisy's gender presentation might have been as complex and personal as that of the Marquise-Marquis de Banneville, whose own journey in the *Histoire* explored themes of identity beyond mere disguise.

In a separate passage, Choisy writes to the actress Montfleury, hinting at a kinship with the titular Marquise-Marquis⁵⁸ de Banneville. "I have no doubt, madame, that the story of the Marquise de Banneville has pleased you," he wrote, confessing his satisfaction at seeing his own life reflected, however loosely, in such an "amiable person." He goes on to acknowledge that, unlike the Marquise, his "prodigious beauty" had not quite offered the same freedoms. This tongue-in-cheek self-identification suggests that Choisy saw elements of his own life and desires mirrored in the Marquise-Marquis character, even if he had not directly authored the work.

The question remains, however, whether Choisy's references to the *Histoire* imply authorship or simply a kind of personal connection to its themes. Some scholars have argued that his familiarity with the tale was suggestive of insider knowledge—indeed, Choisy's memoirs are the only surviving documentation among the potential authors directly referencing the *Histoire*. Yet this fact alone may be coincidental, considering neither Perrault nor Mme L'Héritier left memoirs behind.

⁵⁷ Scott, "Authenticity and Textual Transvestism in the Memoirs of the Abbe de Choisy."

⁵⁸ Choisy, "Aventures de l'abbé de Choisy habillé en femme," 67.

The debate over Choisy's potential authorship of the *Histoire* thus continues, fueled by tantalizing hints in his memoirs. Whether or not he co-wrote or merely inspired the tale remains unresolved, but his open references to the story and self-association with its protagonist have left a lasting imprint on the scholarly interpretation of the *Histoire*. By embedding hints of his own life in these stories, Choisy cultivated a persona that transcends strict biography, reflecting a blend of fiction, performance, and personal truth—a dynamic that still intrigues readers and researchers alike.

◆◆◆ Collaboration between Perrault and Choisy

Jeanne Roche-Mazon's early 20th-century paper, *Une Collaboration inattendue au XVII^e siècle*, remains one of the foundational scholarly works investigating the potential dual authorship of the *Histoire de la Marquise-Marquis de Banneville*. Published in the *Mercure de France* in 1928 alongside a reprint of the *Histoire*'s first edition, Roche-Mazon's article laid out a compelling argument suggesting that Charles Perrault and the abbé de Choisy may have collaborated on this singular fairy tale.

Roche-Mazon bases her theory on the fact that Perrault and Choisy moved within the same intellectual circles and likely interacted both socially and professionally. Both men were part of the Académie française, the pinnacle of literary society in France, and regularly contributed to Paris's literary scene. In fact, Perrault had his famed poem *Le Siècle de Louis le Grand* read at the 1687 Académie reception celebrating Choisy's induction, a moment which reanimated the long-standing Quarrel of the Ancients and Moderns. This debate saw Perrault champion the "Moderns," arguing for the superiority of contemporary French culture over classical antiquity. Choisy, while less publicly vocal on the matter, was closely associated with these ideas and other "Moderns" like Perrault.

Roche-Mazon also notes that Perrault and Choisy were members of a private intellectual group called the Académie du Luxembourg, which met in Choisy's home in 1692. This group served as a literary salon where ideas were exchanged, books were discussed, and members engaged in both intellectual and informal dialogue. Roche-Mazon claims this connection was so deep that Perrault shared private details with Choisy, as Choisy recorded in his *Memoirs to Serve the History of Louis XIV*: "Choisy, in his *Memoirs*, reports a good number of little-known

particularities, on the trial of Fouquet for example, which he tells us he had learned from Perrault.”

The theory that Choisy might have participated in writing the *Histoire* but concealed his contribution is plausible, especially considering the shift in Choisy’s public persona during the late 17th century. Although known for his memoirs documenting his cross-dressing escapades as “Madame de Sancy,” Choisy adopted a pious, respectable image in his later years, presenting himself as an upstanding cleric who had put his more colorful behavior behind him. This reformation may have influenced his reluctance to attach his name to such a fantastical and unconventional work. Roche-Mazon suggests that Choisy may have inspired the storyline or contributed ideas, particularly the gender-bending aspects that bear a strong resemblance to his own life. However, Roche-Mazon is skeptical that Choisy could have penned the more innocent, naïve qualities of Mariane’s character, as these elements contrast Choisy’s signature irony and wit. Instead, she argues that Perrault, known for crafting stories accessible to both children and adults, might have had a greater hand in shaping the story’s tone and style.

Roche-Mazon also brings Madame de Lambert, another key figure in literary circles, into her discussion of the fairy tale’s origins. Known for her influential salon, Madame de Lambert fostered a social environment in which fairy tales gained popularity as both entertainment and moral exploration. Lambert had inspired Perrault’s *Peau d’Ane*, suggesting that her literary circle may have also spurred Perrault and Choisy to experiment with the themes in the *Histoire*. Lambert encouraged Choisy to write his memoirs, which he dedicated to her, implying that she was a significant influence on his literary output. Roche-Mazon implies that Lambert’s guidance could have given both Perrault and Choisy the motivation to experiment with tales that challenged traditional norms of gender and identity in ways both lighthearted and insightful.

There is also the possibility, however, that Choisy’s memoirs themselves borrowed ideas from the *Histoire*, rather than the reverse. Much of Choisy’s memoirs lack precise dates, leaving room for speculation about the chronology of his anecdotes. For instance, Choisy recounts an incident in which he hosted Mademoiselle Dupuis and the parish priest of Saint-Médard at his home, engaging in lively conversation about the *Mercure galant*’s “latest story” of a young man dressing and living as a woman. This story closely resembles the plot of the *Histoire* and Choisy’s real life, but the timing is questionable since the *Histoire* was published in 1695, more than three decades after his adventures as “Madame de Sancy.” Roche-Mazon and later scholars

speculate that Choisy may have fabricated or romanticized parts of this memory to blur the line between life and literature.

Still, Roche-Mazon's research remains the most comprehensive argument for Choisy and Perrault's collaboration on the *Histoire*. Her argument is bolstered by her awareness of the close-knit network between writers, aristocrats, and intellectuals in late 17th-century France. Perrault and Choisy were positioned at the nexus of this network, exchanging ideas, learning from their peers, and perhaps engaging in joint projects that, like the *Histoire*, allowed them to play with ideas of transformation, social mobility, and gender fluidity. As Roche-Mazon points out, Choisy's life experiences and personality align closely with the themes of the *Histoire*, while Perrault's writing style complements the narrative tone. This interweaving of social, intellectual, and personal contexts makes a joint authorship plausible and highlights how much these intellectual friendships influenced literature in early modern France.

❖ ❖ Charles Perrault

Charles Perrault⁵⁹, born on January 12, 1628, in Paris, France, was the youngest of five sons born to Pierre Perrault, a member of the Paris Parliament, and Pâquette Leclerc Perrault. Although initially enrolled at the Collège de Beauvais, Perrault left formal schooling early, choosing instead to pursue an independent study of poetry and philosophy. In 1651, he obtained a law degree and passed the bar, but he soon became disillusioned with the French legal system. By 1654, Perrault shifted his career trajectory, accepting a clerical position under his brother, who was a tax collector.

Perrault's early literary pursuits began in the 1660s, during which time he published a series of poems that garnered attention within the French literary community. His talent drew the interest of Jean-Baptiste Colbert, a prominent aide to King Louis XIV. As a result, Perrault was appointed as an artistic advisor to the royal court. His responsibilities in this role included working for the Office of Royal Buildings, which oversaw cultural and artistic projects. In 1671, Perrault was elected to the Académie française, a prestigious institution that governs matters related to the French language and literature.

⁵⁹ "Charles Perrault | Encyclopedia.Com."

In 1672, Perrault married Marie Guichon and was promoted to Controller of His Majesty's Buildings. After the death of his wife in 1678, Perrault assumed the responsibility of raising and educating their four children, which is likely the source of his interest in writing fairy tales and children's literature.

While Madame d'Aulnoy was the author who first coined the term "conte des fées," Perrault's tales are the ones remembered centuries later. By his death in 1703, Perrault had published a series of novellas and short stories. Many of them were meant to be written accounts of oral stories, and others were contemporary French adaptations of existing stories, like the previously mentioned Puss in Boots or Cinderella.

Today Perrault is considered by scholars to be one of the three authors most likely to have written the *Histoire de la Marquise-Marquis de Banneville*, but for a time he was thought to be the only author. The strongest evidence supporting that premise comes from the first appearance of *La Belle au bois dormant*⁶⁰ in February 1696. It was published in the *Mercure galant*, without an author's name and only an editor's note from Donneau de Visé, which said that "*On doit ce petit Ouvrage à la mesme personne qui a écrit l'histoire de la petite Marquise dont je vous fis part il y a un an...*" ("We owe this little work to the same person who wrote the story of the little Marquise that I told you about a year ago...") implying that both tales were written by the same person.

Since *La Belle au bois dormant* has been a part of the *Mother Goose Tales* from its first publication in 1697, you can see why many assumed that he was indeed the author of the *Histoire de la Marquise-Marquis de Banneville*. Even the preceding note from the original 1695 edition of the *Histoire*, where the author claims to be a woman, is explained by Joan DeJean. In her notes accompanying the MLA publication of the *Histoire*, she makes the case that Perrault was "perhaps the only male writer of his day who would not have minded having his work attributed to a woman."⁶¹ This perspective is supported by Roche-Mazon, who argues that Perrault had a longstanding reputation for supporting women writers and defending female perspectives in literature against contemporary satirists. Perrault openly argued that some of the finest writers of his time were either women or men who, in his view, possessed a similar "feminine" perspective in their writing.

⁶⁰ Perrault, "La Belle au bois dormant," 74.

⁶¹ Perrault, Choisy, and L'Héritier de Villandon, *The story of the Marquise-Marquis de Banneville*, xi

Perrault's defense of female writers extended beyond mere words; he valued the unique qualities women brought to literature and actively resisted the biases of male-dominated literary circles. For instance, Perrault often wrote positively about women's role in shaping the cultural landscape and literary style. His assertion that men could achieve literary greatness by adopting what he saw as a "feminine" way of thinking suggested he valued empathy, emotional depth, and attention to detail—qualities often ascribed to female authors at the time. Such qualities are especially prominent in the *Histoire de la Marquise-Marquis de Banneville*, whose descriptions of Mariane's lavish attire, elaborate jewelry, and grace in social settings reflect a perspective that, to contemporary readers, might seem distinctly feminine. DeJean's assertion finds support in the fact that Perrault, as a schoolboy, honed a keen eye for such details. He translated *On the Apparel of Women* and later authored *Les Fontanges*, a one-act satire focused on the fashion industry. These early works likely informed his ability to vividly describe Mariane's decadent wardrobe in ways that would feel authentic to both male and female readers.

Roche-Mazon also draws notable parallels between the *Histoire* and Perrault's *Cinderella*.⁶² The similarities are particularly evident in the famous ball scene, where both Mariane and Cinderella captivate the crowd with their beauty and elegance. In each story, the protagonist's charm commands the attention of all onlookers, and a prince is drawn to her side in a moment of awe. While in the *Histoire* this prince is not Mariane's love interest, the narrative arc closely resembles the ballroom moments in *Cinderella*, where admiration and social standing play a key role. Such parallels strengthen Roche-Mazon's argument that Perrault likely shaped these elements of the *Histoire*, as he clearly understood the drama and excitement that a grand ball could evoke.

There is additional evidence DeJean supplies in favor of Perrault writing the *Histoire* and allowing it to remain anonymous, which is that there is another case of Perrault publishing without using his own name. This is referring to the first printing of the Mother Goose Tales in 1697 that was published under the name given as "P. Darmancour", hinting at Perrault's 19-year-old son Pierre. While originally taken at face-value; modern scholars believe that Charles Perrault was always the author behind *Contes de ma mère l'Oye*. However, we have no way of knowing if that is due to some sense of playfulness, as DeJean implies, a means to introduce the son to society, or perhaps a way of protecting himself from the critical eyes of the

⁶² Roche-Mazon, "An unexpected collaboration in the 17th century: the abbé de Choisy and Charles Perrault," 518.

“Ancients,” or even written to be the last word in that decade-long literary quarrel on the side of the “Moderns,” suggested by Children's literature scholar Jack Zipes⁶³.

This is referring to the *Querelle des Anciens et des Modernes*⁶⁴ (Quarrel of the Ancients and the Moderns), in which supporters of the literature of Antiquity (the "Ancients") argued against supporters of the literature from the century of Louis XIV (the "Moderns"). Charles Perrault was unquestionably on the side of the Moderns and wrote *Le Siècle de Louis le Grand* (The Century of Louis the Great) and *Parallèle des Anciens et des Modernes* (Parallel between Ancients and Moderns) where he attempted to prove the superiority of the literature of his century, which arguably turned disagreement into something of a literary war. It wasn't always a popular opinion, and his stance on this subject is considered to be the reason why he was forced out of his position as secretary.

Despite these clues, Perrault never publicly claimed authorship of the *Histoire de la Marquise-Marquis de Banneville*. He did not attach his name to the work, nor did he mention it in his known literary discussions. Thus, we are left with an unresolved mystery; while the evidence suggests his influence and hand in its creation, there remains no definitive proof. As a result, Perrault's possible authorship of the *Histoire* remains speculative. Nevertheless, the supposedly “feminine” narrative tone does not detract from the possibility of his having written the *Histoire*. Combined with his advocacy for female writers and recurring thematic elements, this strongly indicate his involvement, if not as the primary author, then at least as a contributor.

◆◆◆ Collaboration between Perrault and L'Heritier

What must be remembered is that the earliest published versions of the *Histoire de la marquise-marquis de Banneville* does not have an author indicated. Visé of the *Mercure galant* seemed to believe it was a woman based on editor's notes, in addition to the claim made by the anonymous author's note itself.

Puis que les Femmes se meslent d'écrire, & se piquent de bel esprit, je ne veux pas demeurer la dernière à signaler mon zele pour mon Sexe.

(Since women are getting involved in writing, and are proud of their wit, I do not want to be the last to show my zeal for my sex).⁶⁵

⁶³ Zipes, *The Oxford Companion to Fairy Tales*, 379.

⁶⁴ Taylor, “The Quarrel of the Ancients and Moderns.”

⁶⁵ “*Histoire de la Marquise Marquis de Banneville*” (1st ed), 14.

Some have long thought that Marie-Jeanne L'Héritier, Charles Perrault's niece, may be the writer. L'Héritier, known for her literary talents and celebrated within *Mercure galant* itself, was a close collaborator with her uncle; they both published fairy tales and prose works in this period, which suggests that her authorship—or a collaborative effort with Perrault—would fit naturally within their shared professional and familial circles.

Scholars have found compelling reasons to support L'Héritier's involvement in the tale. Zipes⁶⁶, in the *Oxford Companion to Fairy Tales*, for example, proposes that the *Histoire* was the product of a collaboration between Perrault and L'Héritier, presumably due to the strong stylistic and thematic overlaps between their other works and the *Histoire*. Paul Delarue similarly suggests⁶⁷ a joint authorship, emphasizing the alignment of the tale's literary style with both Perrault's fairy tale tradition and L'Héritier's narrative voice. L'Héritier's independent work often celebrated themes of wit, social decorum, and female agency, all of which are echoed in the *Histoire de la Marquise-Marquis de Banneville* in the characters of Mariane and the Marquis de Bercourt, who navigate through both male and female spaces.

Even Jeanne Roche-Mazon, who famously argued for Choisy and Perrault's joint authorship, at least had some awareness of the possibility of L'Héritier's influence.

*On serait presque tenté de se demander si l'Histoire de la Marquise-Marquis de Banneville et Marmoisan, qui sont inspirés par des situations exactement inverses et semblent se faire pendant l'un à l'autre, n'auraient pas été composés concurremment par les deux cousins à la suite de quelque gageure amicale...*⁶⁸

(One would almost be tempted to wonder if the History of the Marquise-Marquis de Banneville and Marmoisan, which are inspired by exactly opposite situations and seem to be counterparts to each other, were not composed concurrently by the two cousins following some friendly bet...).

She even speculated that the structure of the *Histoire de la Marquise-Marquis de Banneville* bears an intriguing relationship to *Marmoisan*, another tale associated with L'Héritier. Roche-Mazon mused that the *Histoire* and *Marmoisan* might have been written simultaneously as part of a playful bet or literary exercise between Perrault and his niece, given the way the two tales explore complementary yet contrasting gender identities and transformations. In the *Histoire*, the protagonist Mariane embodies a male disguise for survival and social access, while

⁶⁶ Zipes, *The Oxford Companion to Fairy Tales*, 94.

⁶⁷ Delarue, *Le conte populaire français*.

⁶⁸ Roche-Mazon, "An unexpected collaboration in the 17th century: the abbé de Choisy and Charles Perrault," 518.

Marmoisan features an opposite dynamic, with characters experiencing gender transformation and disguise under different circumstances.

The possibility of such a collaboration speaks to the broader culture of literary salons, where playful wagers, bets, and thematic challenges were common ways for writers to exercise their creativity. This context, combined with Perrault's known investment in promoting the contributions of women writers, supports the idea that he might have collaborated with his niece L'Héritier as a way of encouraging her talents while also contributing to the narrative himself. Perrault's defense of women's literary contributions may have made him more open to attributing authorship to L'Héritier alone, in line with his view that feminine perspectives enriched literature.

Ultimately, the idea of L'Héritier as either sole author or collaborator has compelling support, both in the historical record of *Mercure galant*'s publications and in the stylistic evidence present in the *Histoire*. While definitive proof remains elusive, the story's narrative tone and nuanced engagement with gender roles, social mobility, and identity, strongly suggest the touch of an author who, like L'Héritier, was interested in navigating these themes within the fairy tale genre. This combination of literary insight and social perspective makes a case for her hand in shaping one of the period's most intriguing works.

❖ ❖ Madame L'Héritier

Marie-Jeanne L'Héritier de Villandon was a scholar who surrounded herself with other scholars. Born in 1664 in Paris, L'Héritier was raised in a scholarly environment; her father was a former musketeer who later served as the king's historiographer, and she maintained a close bond with her mother's first cousin, Charles Perrault, whom she affectionately called "uncle." L'Héritier was a member of literary circles that included influential figures like Madeleine de Scudéry, Marie-Catherine d'Aulnoy—who is credited with coining the term *conte des fées* (fairy tales)—and Henriette-Julie de Murat. These women were leaders in the *préciosité*⁶⁹ movement, a literary style that celebrated wit, refined language, and complex social interactions. Through these connections, L'Héritier cultivated her own literary identity, championing women's writing and exploring innovative themes that included fairy tales, poetry, and social critique.

⁶⁹ Raynard, *La seconde préciosité*, 63.

She was received⁷⁰ into *L'Académie des Jeux floraux* of Toulouse in 1696 and the following year, into that of the Académie des Ricovrati of Padua, a prestigious intellectual society. Remarkably, she achieved these honors by the age of thirty-two, an age when most women writers of her time struggled for legitimacy. Unlike many women of her era, L'Héritier never married, choosing instead to support herself through her writing and the patronage of influential figures like the Duchess of Longueville and the Duchess of Épernon. When Scudéry passed away in 1701, L'Héritier even inherited her literary salon, becoming a key figure in the Parisian literary scene. L'Héritier's advocacy for women's literary contributions shaped much of her work. Her first publications included eulogies dedicated to female writers, followed by editorial work on the memoirs of the Duchesse de Nemours, indicating her commitment to preserving and celebrating women's voices.

She wrote poetry, collections of short stories, and even translated Latin works, including Ovid's *Metamorphoses*, into French. Her first collection⁷¹ of four fairy tales—*L'innocente tromperie*, *L'avare puny*, *Les enchantemens de l'éloquence*, *Les aventures de Finette : nouvelles et autres ouvrages en vers et en prose*—were actually published before her uncle, Perrault, published his own Mother Goose Tales. Her frequent contributions⁷² to *Mercure galant*, where her poems won literary prizes and gained royal praise, made her a respected name in French literature. L'Héritier's connection with *Mercure galant* has led some scholars⁷³ to believe that she likely knew the editor, Donneau de Visé, well—an association that further links her to the *Histoire de la marquise-marquis de Banneville*, which was published in that same periodical.

L'Héritier's potential authorship of the *Histoire de la marquise-marquis de Banneville* goes beyond her family ties to Perrault. Scholars suggest that she may have either authored the story alone or collaborated with Perrault. By the time the *Histoire* was published, L'Héritier had already written stories that explored themes of cross-dressing and gender disguise, establishing a precedent that aligns with the themes in the *Histoire*. Her tale *Marmoisan*—a retelling⁷⁴ of “Costanza-Costanzo” from Giovanni Francesco Straparola's *Le Piacevoli Notti*—deals explicitly with cross-dressing and disguises, drawing a strong parallel to the *Histoire*.

⁷⁰ Raynard, *La seconde préciosité*, 63.

⁷¹ L'Héritier de Villandon, *Oeuvres Meslées*.

⁷² “BnF - Portrait de Mademoiselle L'Héritier.”

⁷³ Perrault, Choisy, and L'Héritier de Villandon, *The story of the Marquise-Marquis de Banneville*, xxiv.

⁷⁴ Steinberger, *Women's Stories in “Le Mercure Galant” (1672-1710)*.

Marmoisan tells the story⁷⁵ of a king who, needing soldiers for battle, calls upon his vassals to send reinforcements. The count of Solac, too ill to serve, intends to send his son Marmoisan. However, when Marmoisan dies, the count's daughter Leonore, taking on her brother's identity, decides to join the army in his place. Accompanied by her younger sister Iolande, who disguises herself as Leonore's page, Leonore enters the battlefield and becomes embroiled in a world of deception, chivalry, and political intrigue. As the kingdom falls under threat, the disguised Marmoisan proves herself by saving the prince in battle, ultimately revealing her true identity in a dramatic undressing scene that leads to her marriage with the prince.

Like the *Histoire*, *Marmoisan* exhibits the hallmarks of a fairy tale—royal figures, mistaken identities, a climactic reveal, and a happy ending—without involving any overt magical elements. Both stories feature protagonists who adopt disguises and navigate society from within a different gender role. Yet, a key distinction between the two tales lies in the characters' choices regarding gender identity. While Leonore ultimately returns to her female identity and marries a prince, Mariane in the *Histoire* chooses to live as a woman even after the need for her disguise has passed. This choice makes the *Histoire* particularly progressive for its time, as it subtly advocates for personal autonomy in gender expression—a theme that resonates with L'Héritier's forward-thinking ideals.

In addition to her innovative narration, L'Héritier was a strong advocate for women's literary contributions and intellect. Her feminist ideals were especially evident in her defense of the intellectual abilities of women. She argued that women possessed an innate capacity for wit and sensitivity that could rival, if not surpass, that of men. Perrault, her uncle, shared a similar view and was one of the few male writers of the time to publicly support female authors, claiming that many of the finest writers were either women or men who had cultivated a so-called feminine sensibility. Some scholars, like Joan Dejean, suggest that if the *Histoire* was written by L'Héritier, it could have served as a subtle assertion of a woman's right to navigate society in unconventional ways, just as Mariane navigated her world in the *Histoire*.

The question of whether the *Histoire* was written by L'Héritier alone, or in collaboration with Perrault, remains unresolved. Yet, her close relationship with Perrault, her established literary career, and her focus on women's narratives strongly suggest that she could indeed have

⁷⁵ L'Héritier de Villandon, *Oeuvres Meslées*, 1.

crafted the tale. Whether she served as sole author or co-author, L'Héritier's contributions to the French literary circles of the time embody the values of sophistication, wit, and an early form of feminist ideology that still engages scholars and readers today.

❖ ❖ Other Configurations

Joan DeJean argues that the *Histoire de la marquise-marquis de Banneville* was likely a collaborative effort between all three of the authors previously mentioned. According to her research presented in the MLA edition of the *Histoire*, Charles Perrault, Marie-Jeanne L'Héritier, and the abbé de Choisy each contributed distinct elements reflective of their unique experiences and literary styles. Choisy's involvement may have provided autobiographical nuances to Mariane's background, especially given his experiences with cross-dressing and the fluidity of identity, which likely lent authenticity and depth to the narrative. By drawing on elements from his own life, Choisy may have infused the story—specifically Mariane, Prince Sionad, and the Marquis de Bercourt—with the knowledge of living within two gendered worlds, creating an intriguing foundation that emphasizes the theme of identity beyond disguise. DeJean suggests that Choisy's influence in particular added a layer of self-awareness and societal insight that would not have been present otherwise.

In DeJean's interpretation, Perrault and L'Héritier likely took an active role in the story's development, passing drafts back and forth to refine the plot and embellish certain details. Perrault, with his reputation as a master storyteller and his involvement in the fairy tale tradition, may have helped structure the tale's progression and enrich its tone, balancing Choisy's more personal insights with broader, fantastical elements. DeJean proposes that Perrault's skills in crafting scenes of enchantment and intrigue brought cohesion to the story, while his political astuteness guided its sensitivity to social nuances. His hand may have been responsible for the narrative's structure, including scenes like the ball, which draw on conventions Perrault would establish in other works, such as *Cinderella*. Perrault's experience in navigating the delicate boundaries of courtly life likely informed the way the tale was shaped to avoid offending powerful figures.

L'Héritier's role, according to DeJean, may have been equally significant, as her background in *préciosité* and her passion for women's intellectual and creative voices likely

influenced the story's themes of gender, agency, and wit; qualities that aligned well with the ideals of her own social circle, which supported women's authorship and autonomy. Furthermore, L'Héritier's own tale, *Marmoisan*, featured themes of cross-dressing, gender roles, and disguise—elements she may have woven into the *Histoire* to parallel her previous work, adding richness and continuity to the genre.

Although arguing in favor of the *Histoire* being only written by Perrault and Choisy, Roche-Mazon points out the addition of a foreign background for the character Prince Sionad, an element likely introduced to obscure the real-life connections between him and the Duke of Orléans. In Roche-Mazon's view, this addition helped mask any controversial associations with court figures, particularly since Orléans himself had occasionally defied traditional gender norms, attending some events in women's attire. By giving Sionad a foreign identity, the authors crafted a subtle disguise that preserved the character's essential traits—his beauty and elegance—while avoiding direct reference to Orléans. The care not to avoid political drama is one that would have been necessary for all three authors. It's a detail that exemplifies the careful balance Perrault and Choisy likely maintained, drawing inspiration from recognizable figures while protecting the story and its authors from potential repercussions.

Ultimately, DeJean's theory of a collaborative effort among Perrault, L'Héritier, and Choisy suggests that the *Histoire de la marquise-marquis de Banneville* was not only a work of literary imagination but also a carefully constructed narrative reflective of its authors' experiences, ideals, and the intricate social landscape of the time. By integrating Choisy's firsthand experience with gender identity, Perrault's storytelling mastery, and L'Héritier's feminist perspective, the tale becomes a multifaceted story that touches on themes of identity, social perception, and personal freedom—topics that remain relevant in literature and society to this day.

The influential role of the editor, Jean Donneau de Visé, in shaping public perception of the tale and its authorship is also an essential factor to consider. Visé was not merely a gatekeeper of content; he was a shaper of its public reception. The anonymous or pseudonymous nature of many stories in *Le Mercure galant*, combined with Visé's editorial choices, allowed him to subtly guide readers' interpretations. For example, his introductory note to the *Histoire* and, later, to *La Belle au bois dormant*, implied a female authorship, likely to appeal to readers' appetite for novelty and to intrigue those aware of the period's increasing number of

female-authored literary works. By presenting the *Histoire* as a tale written by a woman, Visé's decision to allow the tale to be presented without an identified author may also reflect an intentional ambiguity that catered to his audience's taste for mystery and speculation. His editorial framing invited readers to wonder about the author's identity, contributing to a lasting aura of intrigue around the story. In the absence of concrete attribution, Visé effectively positioned the *Histoire* as a tale that could embody the progressive and playful voice of any woman in the literary community, rather than tying it to a specific individual. This ambiguity opened up a creative space that encouraged interpretations beyond the circles of Perrault, L'Héritier, and Choisy, expanding the literary imagination of readers who may have seen in the *Histoire* an anonymous woman's challenge to gender roles.

However, the assumption that the writer of the *Histoire de la Marquise-Marquis de Banneville* MUST be one of the three authors above, or some combination, does a disservice to the wealth of writers—many of them women—who submitted works to Le Mercure Galant. It narrows the scope of potential authorship by focusing on prominent figures and those with identifiable literary connections to the tale. This assumption limits our understanding of the diverse pool of contributors, many of them women, who published anonymously in the *Mercure galant*. The popular periodical, known for its inclusivity, served as an important platform for women writers whose works often explored themes considered unconventional or socially challenging for the time. As Mary Elizabeth Storer points out, assuming that the *Histoire de la Marquise-Marquis de Banneville* must be the work of established male or closely connected female authors underestimates the scope of anonymous women⁷⁶ writers who also had the creative capacity to produce such a nuanced and subversive tale.

Women authors in the late 17th century, particularly those associated with salons, were known to submit work to the *Mercure galant*, and welcomed by Visé. It's possible that the author of the *Histoire* was one of these lesser-known or anonymous women who, lacking the privilege or connections of a Perrault or a L'Héritier, was nonetheless able to utilize the space offered by the *Mercure galant* to explore progressive themes. This inclusivity of women authors meant that the periodical became a testing ground for innovative storytelling, and the *Histoire* fits within this tradition of literary experimentation.

⁷⁶ Storer, *Un épisode littéraire de la fin du 17e siècle*.

Ultimately, both Storer's advocacy for an anonymous female author and the editorial influence of Visé remind us of the expansive network of contributors to *Le Mercure galant*. This was a period when women increasingly found literary freedom through salon culture and periodicals, using their voices to explore themes of identity, disguise, and personal agency—qualities that resonate strongly within the *Histoire*. By viewing the tale as potentially emerging from this broader, diverse literary community, we recognize it as part of a collective push toward modern ideas, rather than the isolated product of elite literary figures alone. This broader perspective not only honors the innovative voices of early women authors but also restores the *Histoire* to its place within a vibrant, collaborative cultural tapestry that included and supported a multitude of voices.

◆ Whodunit? A brief history of authorship attribution

Was it the butler who did it, in the lounge with the notebook? Or was it the niece in the study with the typewriter? Perhaps it was the cook in the kitchen with the tablet...

The question of authorship can feel a bit like a literary mystery, one that requires careful detective work to solve. The history of the written word, however, tells us that for most of human history, we weren't so concerned with solving that mystery. For thousands of years, it wasn't the person behind the words that mattered, but the words themselves. The ancient world was full of anonymous writings, and the absence of an author's name wasn't a problem; it was often the expected norm. Clay tablets from Mesopotamia detailing complaints or receipts⁷⁷ didn't include the signature of the scribe. Greek philosophical scrolls were circulated without the name of the person who copied them, because the focus was on the ideas they contained, not the person who happened to put ink to papyrus. Even religious texts like the Bible passed through many hands over the centuries, each copyist contributing without expecting personal recognition. Anonymous⁷⁸ has a wide and varied catalog of work.

Only in more recent centuries has authorship become something we care about deeply, particularly as intellectual property and the idea of "the author" as a celebrated individual became more important. Today, authorship is central to how we understand literature, copyright,

⁷⁷ Magazine and Anderson, "Researchers Decipher Cuneiform Tablet—and Discover It's a Furniture Receipt."

⁷⁸ From the Ancient Greek ἀνόνημος (anónimos, "without name")

and even historical legacy. But while the modern literary world often demands a clear attribution for a text, many works remain anonymous or pseudonymous. In some cases, this is a deliberate choice, as seen in the wide and varied catalog of work by “Anonymous” today. For others, it’s a mystery left to scholars, historians, and increasingly, computers to solve.

Determining authorship is often based on two key elements: evidence and analysis. In the simplest cases, authorship can be confirmed through clear, external evidence. A book might bear the name of its author on the title page, or a letter might be signed at the bottom. This kind of direct evidence is the most straightforward way to attribute a work. However, in the absence of such a signature, other types of evidence must be collected and pieced together. Contextual clues can offer insights into who may have written a particular text. Take, for instance, an anonymous social media post made by a student at Thesis University. It mentions being a member of both the debate team and the hockey team; a commonality shared by only three students at the school. They also mention a study abroad trip to Italy. In this instance there is only one student that debates and plays hockey and studies Italian. From this data we can surmise that this one student was the author of the anonymous post.

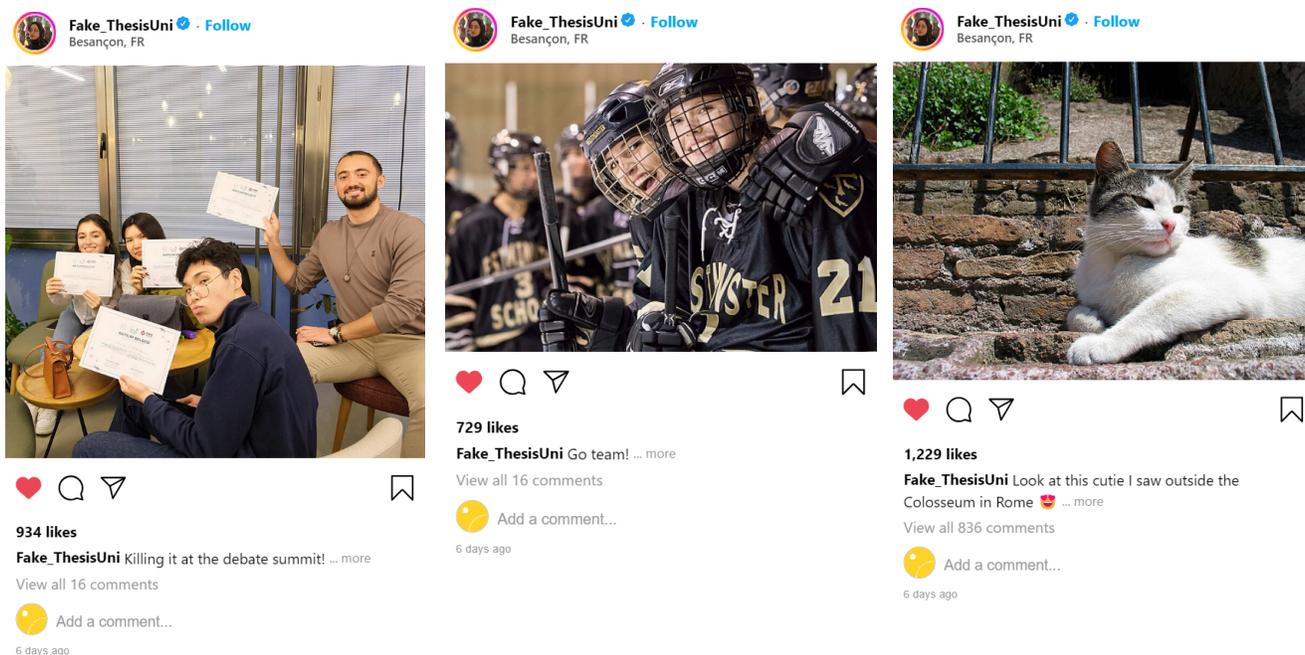


Figure 3. Instagram template generated by Zeob, with photos sourced from Wikimedia Commons.

In many cases, however, authorship is not so conclusive. Lack of contextual clues—or conflicting ones—can make attribution difficult. Historical documents, literary works, and other writings often come with little or no external context, making the task of attribution much more difficult. This is where quantitative analysis comes into play, using a more scientific approach to authorship determination.

One of the earliest examples of such analysis can be traced back to the study of Shakespeare's works. For centuries, scholars have debated whether Shakespeare was the sole author of the plays attributed to him, or if other playwrights⁷⁹ contributed. Early analysis focused on patterns of meter, rhyme, and word endings⁸⁰, searching for stylistic inconsistencies that might point to multiple authors. Today, these methods have evolved into more sophisticated techniques such as stylometry, the statistical study of writing style. By examining word frequency, sentence structure, and other linguistic features, stylometry allows us to compare a questionable text to a known body of work and determine whether the two share a common author.

The rise of computational methods has revolutionized this process. Modern computers can analyze massive datasets in a fraction of the time it would take a human, comparing texts across a wide range of stylistic markers. In practice, this methodology is applied through stylometric analysis by way of Natural Language Processing—a technique used to identify an author's distinct writing style by examining elements like word frequency and sentence structure. By comparing the *Histoire* with other works, I sought to establish statistical correlations that could provide insight into the true authorship of the fairy tale.

Stylometry is the study of writing style, seeking to identify and quantify distinctive patterns, traits, and characteristics within texts. It's a multidisciplinary field that intersects linguistics and computational analysis when being used to determine authorship. Preparing data for stylometry is a crucial step in extracting meaningful insights from texts. It involves transforming raw text into structured and analyzable formats, enabling the application of computational techniques. The quality and diversity of the data significantly impact the results and the generalizability of findings. For authorship attribution tasks, a selection of texts of known works from potential authors is compiled—in this case, texts from all three potential authors; de Choisy, Perrault, and L'Heritier.

⁷⁹ Malone, *A Dissertation on the Three Parts of King Henry VI. Tending to Shew That Those Plays Were Not Written Originally by Shakspeare.*

⁸⁰ Kahan, "I Tell You What Mine Author Says," 830.

Computational Analysis

The introduction of computational analysis to the scholarly field of the humanities has transformed how researchers study literature, history, and culture, but it remains a point of contention. Traditional humanities research emphasizes close reading, interpretation, and qualitative analysis; methods that have long been central to the field. However, computational tools like text mining, Natural Language Processing (NLP), and stylometry offer a new approach, allowing scholars to analyze vast amounts of data quickly and at a scale that would be impossible through traditional methods.

One key advantage of computational analysis is its ability to detect patterns, trends, and anomalies across large bodies of texts. For example, these tools can identify shifts in language use, stylistic elements, or cultural themes over time, which may not be noticeable through close reading alone. This opens up new possibilities for cross-comparative studies between different periods, genres, and authors, enriching the scope of humanities research. Moreover, computational methods facilitate large-scale projects by enabling scholars to analyze entire digital archives, making it easier to uncover hidden connections and broader insights.

However, the rise of digital humanities has sparked debates⁸¹ among scholars. Traditionalists argue that close reading and human interpretation remain irreplaceable, fearing that computational analysis may oversimplify complex texts. Meanwhile, digital humanities scholars advocate for the integration of both approaches, suggesting that modern tools can complement, rather than replace, traditional methods. Despite the contention, computational analysis is gaining traction, offering a powerful way to expand the depth and breadth of humanities research.

The debate between traditional and computational methods in the humanities may soon become a moot point due to the ubiquitous presence of computers, smartphones, and digital tools in the lives of modern students and researchers. Today's students, raised in a digital world, are growing increasingly more accustomed to using technology for research, writing, and communication. This familiarity with digital platforms makes it easier to integrate computational methods into their academic work. As a result, many emerging scholars are not only comfortable

⁸¹ Greetham, "The Resistance to Digital Humanities."

with but *expect* to use modern tools like text mining software, digital archives, and data visualization programs in their research.

This generational shift means that the divide between "old" and "new" methods will likely diminish over time as digital approaches become a standard part of humanities education. Rather than replacing traditional methods, computational tools are more likely to be seen as essential supplements that can deepen textual analysis and broaden research possibilities. In this sense, the future of humanities scholarship may lie in a blended approach, where the close reading of individual texts and the large-scale analysis of patterns coexist, allowing scholars to take advantage of both traditional and digital methods in their pursuit of knowledge.

◆ What is NLP?

Natural Language Processing⁸², or NLP, can sound intimidating, but in practice it's essentially about teaching computers to interact with human language—understanding, analyzing, and even generating it. The process starts with selecting the texts that you want to work with. In today's world, many of these texts are "born-digital," meaning they were originally created in digital format rather than being scanned or digitized from physical sources. Born-digital texts can include emails, social media posts, e-books, and websites. These documents are already in machine-readable format, which makes them easier to process compared to older texts that need to be scanned and converted into digital format.

For older, printed, or handwritten texts, a process called Optical Character Recognition⁸³ (OCR) is typically used. OCR converts scanned images of printed text into machine-readable text. This process, however, isn't perfect and may introduce errors, especially if the text is old, damaged, or formatted in an unusual way. In the case of handwritten manuscripts, a more specialized process called Handwritten Text Recognition⁸⁴ (HTR) is employed, which is designed to recognize and convert handwriting into text. While born-digital content has made NLP work much more straightforward, these processes of OCR and HTR remain essential for working with historical documents and manuscripts.

⁸² Greenberg, "The Applicability of Natural Language Processing (NLP) to Archival Properties and Objectives," 401.

⁸³ Dietrich, "OCR vs. HTR or 'What Is AI, Actually?'"

⁸⁴ Ibid.

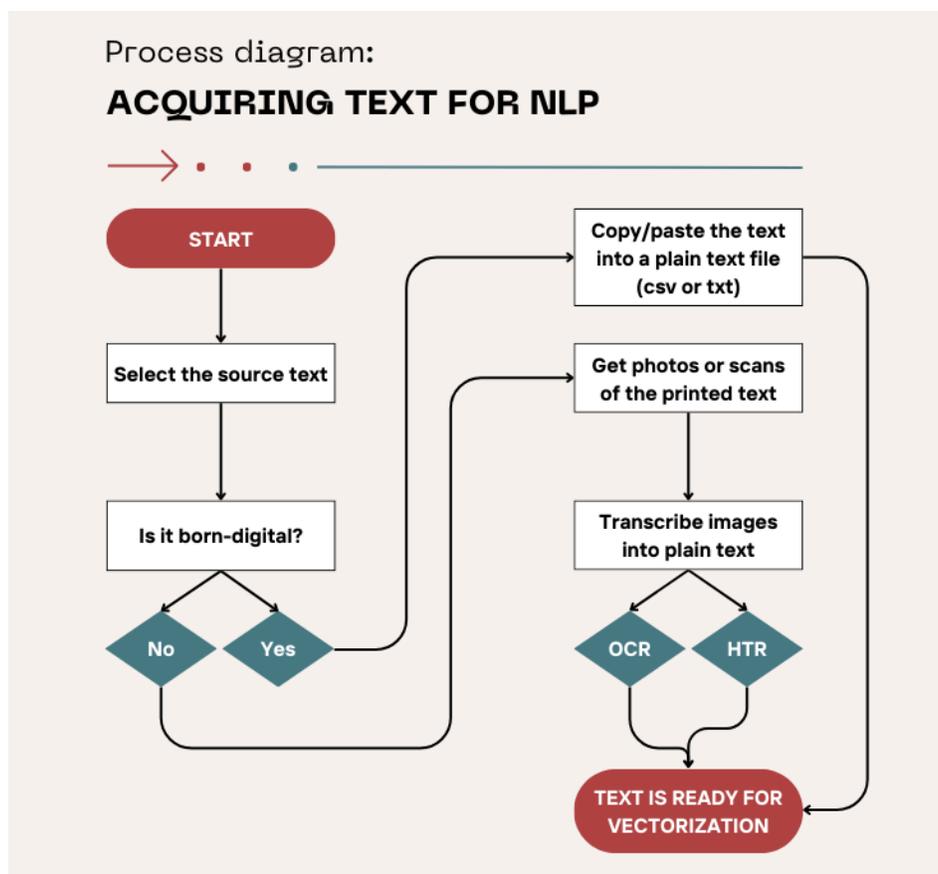


Figure 4. Flowchart created in Canva.

Once the text is available, ideally as a simple TXT or CSV file, the next step is to transform the text into something a machine can understand, because words and letters are meaningless to a computer that only really understands numbers. Transforming the text in this way is referred to as text vectorization; it transforms the text into numerical representations so that it can be analyzed by machine learning models or algorithms. The result turns the *text* into a *dataset* that is now ready to work with.

One common method is to represent each word as a unique number or a set of coordinates in a "vector space." Another is the "Bag of Words"⁸⁵ approach, where the frequency of words is used to build a dataset. More advanced approaches like word embeddings can capture semantic relationships between words by placing them in a vector space where similar words are closer together.

⁸⁵ Jayaswal, "Text Vectorization."

However, before jumping into the analysis, it's important to clean the data. Data cleaning in NLP involves removing irrelevant or potentially confusing information that could skew the results. The ultimate goal of data cleaning and text vectorization is to prepare the text for computational analysis, of which there are many methods and algorithms.

In my case, for authorship attribution—determining who wrote a particular text—common techniques include analyzing word frequencies, punctuation patterns, or sentence structures. Stylometric techniques, like Kilgariff's Chi-Squared method or John Burrows' Delta⁸⁶ method, are often used in such cases to measure stylistic similarities or differences between texts. Machine learning models might be trained to recognize patterns in the writing, enabling more sophisticated and nuanced analysis. However, the downside of machine learning models is that they are very data hungry⁸⁷ and only function well when there is a vast amount of text to work from.

After preparing the text and choosing the appropriate algorithms, the final step is to actually run the analysis. This could involve training machine learning models, comparing different texts, or calculating statistical measures of word usage. The results could reveal hidden patterns, differences between authors, or even shifts in language over time. With this combination of linguistic processing, data cleaning, and computational analysis, NLP opens up a wide range of possibilities for understanding and interpreting texts in new ways.

◆ Methodology

The process of authorship attribution began with selecting and acquiring the texts I planned to use. I started with identifying early versions of the *Histoire de la Marquise-Marquis de Banneville*. I gathered these from the rare book collections at the Bibliothèque nationale de France (BnF) where first and second editions are available as digital copies, and the 1723 edition of the *Histoire* is available in print form; I also collected texts from the three potential authors. Once acquired, the texts were digitized and preprocessed to form a dataset suitable for computational analysis. “Preprocessing” in this case means data cleaning.

⁸⁶ Laramée, “Introduction to Stylometry with Python.”

⁸⁷ Keerio, “Why Deep Learning Is so Called Data-Hungry.”

This included removing special characters, punctuation, or non-standard letters like accented characters. In some cases, all letters might be converted to lowercase to ensure uniformity. When working with older texts that were digitized through OCR, this process also involved correcting transcription errors that could have been introduced during scanning. The goal was to make sure the text is as clean and uniform as possible.

Another key part of the data cleaning was deciding whether or not to remove stop words. Stop words are common words in any language like "the," "is," "and," or "of" that do not carry specific meaning on their own but are necessary for constructing sentences. In some NLP applications, like topic modeling, these stop words might be removed to focus on content-heavy words. In contrast, to some analysts the stop words might carry subtle but important clues about a writer's style and therefore may be left in the dataset.

The next step involved creating a dataset from these texts, using NLP techniques in Python. Python is a versatile and accessible programming language known for its efficient high-level data structures and straightforward approach⁸⁸ to object-oriented programming. Designed to be easy to learn, it has become widely popular across various fields, from web development to data analysis. (Its creator, Guido van Rossum, named the language after the BBC comedy show *Monty Python's Flying Circus*, adding an element of humor to the language's origins.) Python is also classified as FLOSS (Free/Libre and Open Source Software), meaning its source code is freely available⁸⁹ for anyone to use, modify, and distribute.

This open-source philosophy has encouraged a vast community of developers to contribute additional tools and libraries, enhancing Python's capabilities. Many libraries are made available on platforms like GitHub, enabling others to benefit from and expand on existing resources. In the context of Python coding, a GitHub library is a collection of reusable code, tools, and resources shared on GitHub, often organized as a repository. It typically includes Python modules, functions, and classes designed to perform specific tasks or add functionalities; everything from data analysis to machine learning). Developers can integrate these libraries into their projects by cloning or downloading the repository or installing it. Libraries on GitHub often come with documentation, examples, and licensing information to help users understand and use the code effectively.

⁸⁸ C.H., "About Python · A Byte of Python."

⁸⁹ Ibid.

The collaborative nature of Python has produced a variety of specialized tools, including Fast Stylometry, which is just one example of a library resource that can be used for handling complex analyses. I'll be incorporating Fast Stylometry into my project in order to utilize John Burrow's Delta⁹⁰ method. It's a commonly used statistical method for authorship attribution in which the algorithm calculates the "delta" or difference between the frequencies of common words (or other stylistic markers) in an unknown text and in samples of known authors' works. By using standardized differences, Burrows's Delta highlights the distinctiveness of an author's style, helping to match a text to its most similar known author profile.

❖ ❖ Acquiring Text

Using resources like WorldCat.org, I located editions of the *Histoire* and identified comparison texts from the same historical period. When selecting comparison texts for author attribution, I knew it was best to choose works that were similar in time period, genre, and style to the text in question. This ensured a fair basis for comparison, as patterns of language use, themes, and stylistic choices are more likely to align.

The goal was to find works with similar themes, subject matter, or time period to compare linguistic styles. These included Charles Perrault's *Mother Goose Tales*, which are contemporaneous fairy tales, as well as the particular edition of *Sleeping Beauty* that was published in the *Mercure galant* in the same year as the *Histoire's* second edition. I also selected Madame L'Héritier *Oeuvres Meslées*, particularly including *Marmoisan*, for its thematic similarity of cross-dressing, and the *Memoirs of the Abbot de Choisy Dressed as a Woman* by the abbé de Choisy, given its relevance to gender fluidity, as well as a volume of his *Memoires Pour Servir A L'Histoire de Louis XIV* because its description of the royal court reflects experiences had by the *Histoire's* main character.

To obtain these texts, I navigated the Gallica website of the Bibliothèque nationale de France (BnF) for digital versions and used a scan tent to photograph print-only editions, such as the 1723 third edition of the *Histoire*, as well as getting photographs of the microfiche copy of the 1928 La Centaine *Histoire* reprint. Some of the comparison texts I was able to find online via Wikisource and the Internet Archive.

⁹⁰ Laramée, "Introduction to Stylometry with Python."

◆◆◆ Multiple editions of the *Histoire*

At the beginning of my project, I acquired the text of the *Histoire de la Marquise-Marquis de Banneville* from Sorbonne University's *Mercure galant* project⁹¹ created by their Observatoire de la vie littéraire department (licensed under creative commons CC BY-NC-ND 3.0 FR).

Le corpus propose une édition des articles du Mercure galant relatifs à la vie littéraire (critiques, poésies, histoires, actualités éditoriales et académiques), à la musique, aux spectacles et aux cérémonies depuis la fondation du journal (1672) jusqu'en 1710. À la navigation par la table des matières et à la recherche en plein texte s'ajouteront bientôt les planches de musique notée et les gravures littéraires.

The corpus offers an edition of the articles from the Mercure galant relating to literary life (criticism, poetry, stories, editorial and academic news), music, shows and ceremonies from the founding of the newspaper (1672) to 1710. Navigation via the table of contents and full-text search will soon be joined by notated music plates and literary engravings.

Their collection covered both the 1695 first edition of the *Histoire* as well as the second edition that was split across two months in 1696. However, I eventually discovered that whatever transcription process the Sorbonne used, it resulted in pages of the story missing from the final transcription in both editions. To fill the gap, I sought out the versions of the *Histoire* on Gallica, the BnF's digital arm, in order to acquire PDF copies of the *Mercure galant*'s relevant publications. From there I split out the relevant pages. In order to validate the Sorbonne's transcription, I did a side-by-side comparison of the magazine's images and the transcription and manually typed in the missing sections when necessary.

⁹¹ "Mercure Galant | OBVIL."

BnF Gallica

PAR TYPES DE DOCUMENTS PAR THÉMATIQUES PAR AIRES GÉOGRAPHIQUES BLOG

Home > 6 285 Results Page 1 on 419 > 67 Issues found, page 1 on 5 > Consultation

Mercure galant 01 février 1695

ISSUE BY DATE

< 1694 **1695** 1696

Jan. **Feb.** March Apr. May June
Jul. Aug. Sept. Oct. Nov. Dec.

1 2 3 4 5 6
7 8 9 10 11 12 13
14 15 16 17 18 19 20
21 22 23 24 25 26 27
28

See all issues for the year 1695

SYNTHESIS >
A T >

14 MERCURE
HISTOIRE
DE LA MARQUISE,
Marquis de Banneville.
Puis que les Femmes se
mellent d'écrire, & se
piquent de bel esprit, je ne
veux pas demeurer la dernie-
re à signaler mon zele pour
mon Sexe, & il ne tiendra
pas à moy qu'on ne nous
croye de grands personna-
ges, malgré toutes les petites

GALANT. 15
façons dont nous ne scau-
rions nous defaire. En effet,
que lque guindées que nous
foyons dans nos Ouvrages,
on y voit la Femme en mille
endroits, & les grands senti-
mens outrez, forcez, lubli-
mes, ne scauroient cacher
aux yeux du Lecteur attentif
une certaine mollesse, un
certain foible qui nous est
naturel, & où nous retom-
bons toujours. Il ne faut donc
pas nous donner pour plus
que nous ne valons. Croire
qu'une jeune Fille assez jolie,
élevée parmy les rubans, soit

ZOOM Page 14

Downloading and printing

Choose the file format:

PDF

JPEG (only current page)

TXT

Download

full document

selection from the document

By ticking this box, I acknowledge that I have read and understood the [Conditions for use](#) and I accept them.

Download Cancel

Figure 5. Screenshots of Gallica's digital interface.

The third edition of the *Histoire* was a standalone text that was never digitized. As far as I'm aware it was never reprinted either. I consider it a valuable addition to the dataset because it

has two additional scenes. It will be interesting to see how the development of the story over time might impact the results.



Figure 6. Personal photo taken using a BnF ScanTent.

Using a scan tent at the BnF in Paris proved invaluable for photographing rare books, especially those not available in digital form. The scan tent allowed for high-quality, consistent lighting, which is essential for capturing detailed images of delicate, old texts without causing damage. This setup ensured that each page was evenly lit and free from glare or shadows, which can often obscure important details in rare manuscripts. The carefully controlled environment also minimized handling of the fragile books, preserving their condition. These high-resolution images were crucial for running OCR later, as clear and well-lit photos lead to better accuracy in text recognition.



Figure 7. Personal photo of a BnF ScanTent.

In addition to the 1723 edition, I also took photographs of the microfiche version of the 1928 *La Centaine* reprint of the *Histoire*. Ultimately, I decided not to use this text as it proved to be an identical reprint of the first edition, but by checking the microfiche I was able to verify that it was a reprint as well as establishing the editor's note regarding the belief that the *Histoire* was a collaborative attempt between Charles Perrault and the abbé de Choisy.

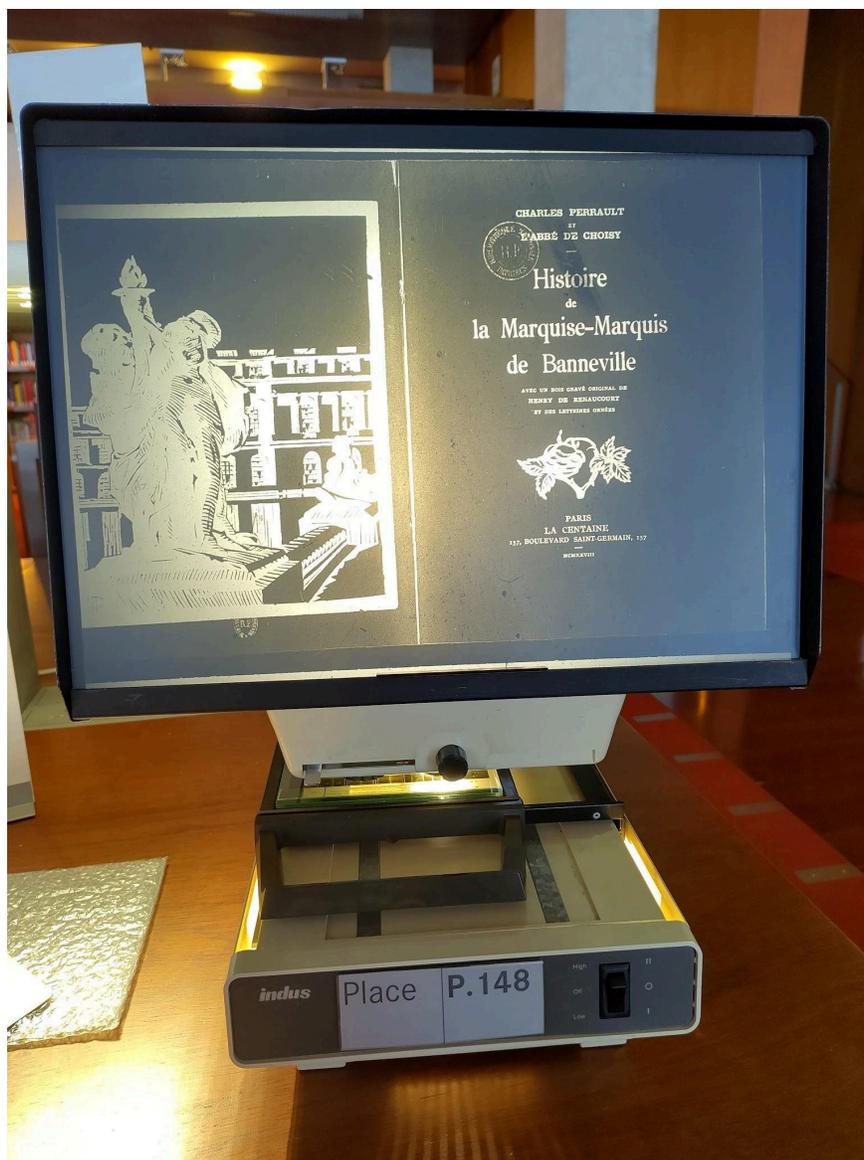


Figure 8. Personal photo of a BnF microfiche station displaying a *Histoire* title page.

◆◆◆ Charles Perrault texts for comparison

I began by selecting Charles Perrault's *Les Contes de ma mère l'Oye* (Mother Goose Tales) for comparison. I chose the first edition, published in 1697 under the title *Histoires ou Contes du temps passé*⁹², because it appeared only a year after the second edition of the *Histoire de la Marquise-Marquis de Banneville*. This proximity in time made the text an ideal candidate for comparison, as it minimized the influence of later editing and refinement that could obscure

⁹² Perrault, "Contes ou Histoires du temps passé - Les Contes de ma Mère l'Oye."

Perrault’s earlier writing style. By focusing on this edition, my goal was to capture the author’s most authentic linguistic patterns of this time period.

The text was sourced from French Wikimedia, which has a complete transcription of this edition. The stories included are *La Belle au bois dormant* (Sleeping Beauty), *Petit Chaperon rouge* (Red Riding Hood), *La Barbe bleüe* (Bluebeard), *Le Maistre Chat, ou le Chat botté* (Puss in Boots), *Les Fées* (The Fairy), *Cendrillon, ou la Petite Pantoufle de verre* (Cinderella), *Riquet à la houppe* (Riquet with the Tuft), and *Le Petit Poucet* (Tom Thumb).

For curiosity’s sake I also be included in the dataset the version of *La Belle au bois dormant* that was published⁹³ in the *Mercure galant* in February 1696. This is the one with the editor’s note that it was written by the same person who wrote the *Histoire*—it will be interesting to see what the comparison results will be.



- [A Mademoiselle](#), par P. Darmancour
- [La Belle au bois dormant](#), conte.
- [Petit Chaperon rouge](#), conte.
- [La Barbe bleüe](#), conte.
- [Le Maistre Chat, ou le Chat botté](#), conte.
- [Les Fées](#), conte.
- [Cendrillon, ou la Petite Pantoufle de verre](#), conte.
- [Riquet à la houppe](#), conte.
- [Le Petit Poucet](#), conte.

Figure 9. Screenshot of Wikimedia “Contes ou Histoires du temps passé - Les Contes de ma Mère l'Oye.”

⁹³ Perrault, “La Belle au bois dormant.”

◆◆◆ Abbé de Choisy texts for comparison

In addition to Perrault, I selected texts by the abbé de Choisy for comparison, even though he did not write traditional fairy tales. Choisy's memoirs, which chronicle his experiences within the French court, interactions with the nobility and time spent dressed as a woman, contain content similar to that of the *Histoire de la Marquise-Marquis de Banneville*. Choisy's memoirs were first published in a partial collection as *Mémoires pour servir à l'histoire de Louis XIV*⁹⁴ in 1727, only three years after he died. A more complete collection, serving as a sequel, was first published in 1736 as *Mémoires de Madame la comtesse des Barres*⁹⁵. This second collection is the one that describes Choisy's upbringing and habits of crossdressing; it was reprinted as *Aventures de l'abbé de Choisy habillé en femme* in 1862.

I was able to source an already-digitized copy of *Aventures* from Project Gutenberg, though this edition was from 1920. To ensure accuracy, I compared this version to the oldest copy available⁹⁶ on Gallica. After a detailed comparison, I verified that the content had not been modified, with the only difference being the sequence of the journal entries. Since sequence order would not affect the types of analysis I intended to perform, I concluded that the 1920 edition was acceptable for my purposes.

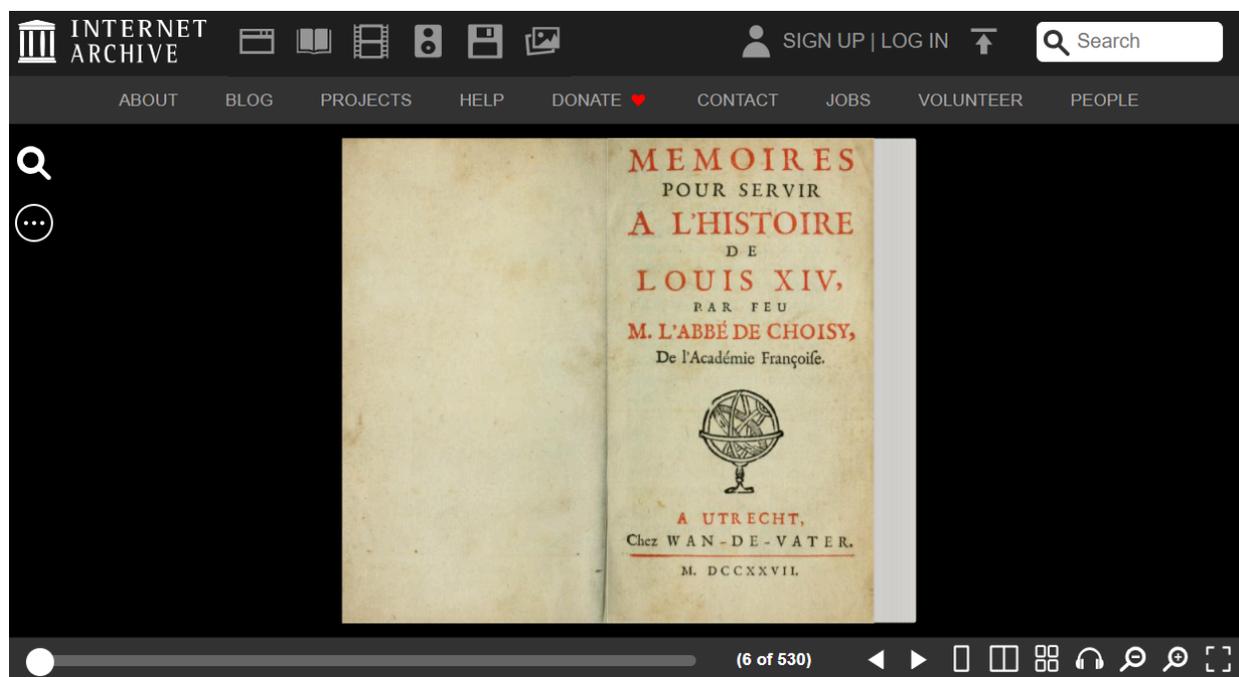
Additionally, I also included his *Mémoires pour servir à l'histoire de Louis XIV* that has similar content to the *Histoire* and was published closer in date to it, which I sourced an early edition from the Internet Archive. I also found a copy of his *Journal du voyage de Siam*⁹⁷ from 1687, because Choisy's memoirs lack specific dates, and I wanted to secure a text that was closer to the first and second editions of the *Histoire*. The inclusion of both memoirs and the journal established linguistic and thematic baselines that will support a comprehensive analysis by situating Choisy's works within the social and literary context of late Seventeenth-century France.

⁹⁴ Choisy, *Mémoires pour servir à l'histoire de Louis XIV*.

⁹⁵ Choisy, *Histoire de Madame La Comtesse Des Barres*.

⁹⁶ Choisy, *Aventures de l'abbé de Choisy Habillé En Femme*.

⁹⁷ Choisy, *Journal Du Voyage de Siam*.



Mémoires pour servir à l'histoire de Louis XIV

by Choisy, abbé de, 1644-1724



Figure 10. Screenshot of the Internet Archive's digital interface.

The text for the *Mémoires pour servir à l'histoire de Louis XIV* was originally drawn from the Internet Archive, which in turn got its text from Google OCR. I found this OCR had resulted in many errors, likely due to Google's algorithms having been training on more recent printing samples and therefore struggling with the antique formatting and lettering of the Seventeenth-century. I would have to seek an alternative method of acquiring the text. (See the Transcription section for more).

◆◆◆ Madame L'Héritier texts for comparison

Another key work in my authorship attribution study that I chose was *Œuvres meslées*, published in 1696⁹⁸ by Madame L'Héritier, the niece of Charles Perrault. Like Perrault, Madame L'Héritier produced works within the genre of fairy tales, making her collection an ideal

⁹⁸ L'Héritier de Villandon, *Oeuvres Meslées*.

comparison for the *Histoire de la Marquise-Marquis de Banneville*. The collection's long title—*Œuvres meslées : contenant L'innocente tromperie, L'avare puny, Les enchantemens de l'éloquence, Les aventures de Finette : nouvelles et autres ouvrages en vers et en prose*—reflects the diversity of her works, which blend verse, prose, and short stories.

I have a particular interest in the collection of *L'innocente tromperie* (The Innocent Deception). Mentioned earlier, that is the story of Marmoisan, a tale featuring themes of gender disguise and fluidity. This genderqueer element aligned with the themes found in *Histoire de la Marquise-Marquis de Banneville*. I sourced a first edition digital copy of this work from Gallica and was able to download a PDF to create a transcription.



Figure 11. A screenshot of Gallica's copy of *Œuvres meslées*.

❖ ❖ Transcription: Transkribus and Tesseract

Much of the text I intended for the analysis dataset was able to be acquired digitally, because previous scholars had already transcribed the works. However, for the remaining works that had not yet been transcribed, or at least not accessible online, I used PDFs converted to PNG images as well as my own digital photos to create machine-readable texts. Working with

Seventeenth-century texts, which often feature ornate typefaces, special characters, and non-standardized spelling, presented some challenges in digital transcription.

For this transcription process, I employed both Transkribus and Tesseract OCR. Transkribus is an advanced transcription and OCR platform that supports scholars in transforming historical documents into searchable and editable formats. It uses Optical Character Recognition (OCR) for printed text and Handwritten Text Recognition (HTR) for manuscript sources. In Transkribus, users can also train custom models tailored to specific fonts or handwriting styles, increasing transcription accuracy significantly, especially for complex or unique historical typefaces. For this project, I relied on Transkribus's "Print M1" model⁹⁹, which offers a Character Error Rate (CER) of 2.20%, and is meant for multiple languages, including French. The fact that Transkribus has a user interface, rather than requiring you to work directly with code, and automated text region selection and recognition makes it more accessible to a lot of scholars.

Transkribus, in particular, was my initial tool, given my familiarity with it from an internship with the Paris Bible Project in 2023, where I learned how to leverage its capabilities for historical texts. Of the texts that I needed for my dataset, I used Transkribus on L'Héritier's *Œuvres meslées*, and extracted *Marmoisan* from that collection. I also employed Transkribus towards Choisy's *Journal du voyage de Siam*.

⁹⁹ "Transkribus Print Multi-Language."

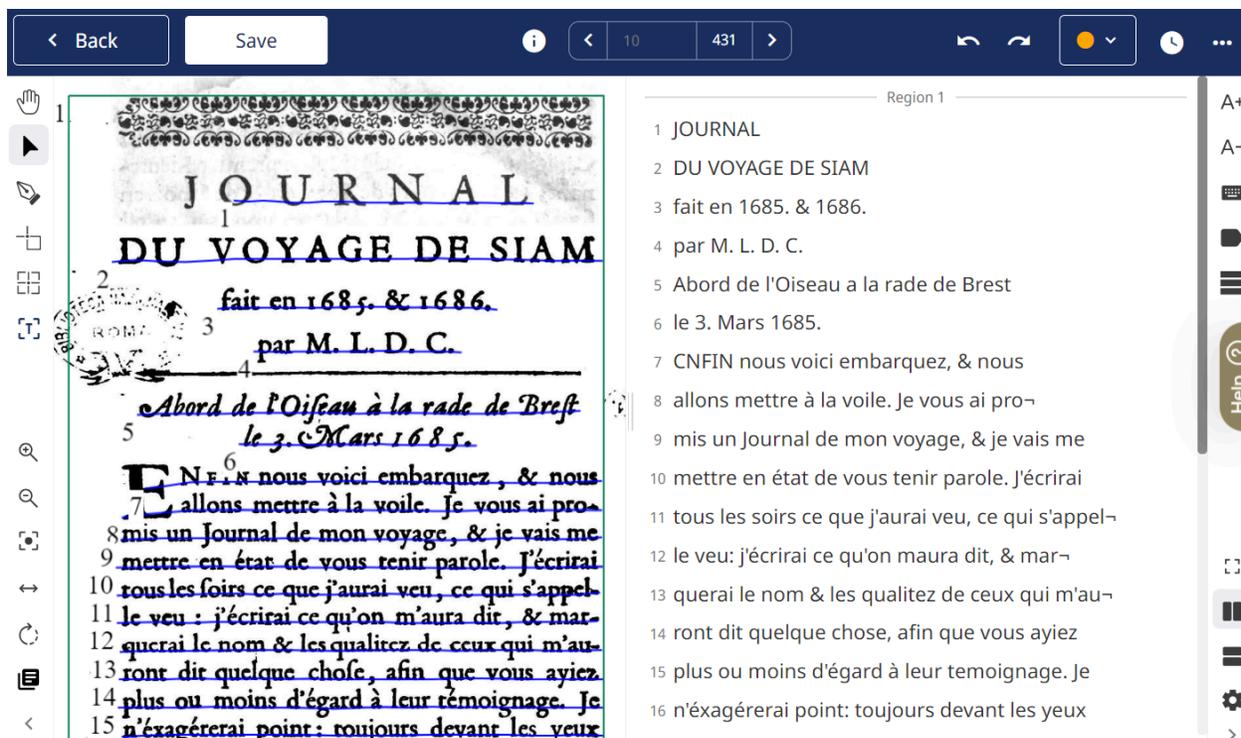


Figure 12. Screenshot of the Transkribus interface.

As I progressed through the course of my master's degree and a second internship, this one with the University of Ghent and their Center for Digital Humanities, I gained more familiarity with using Python. Once I acquired the photos of the *Histoire's* third edition using the ScanTent at the BnF, I decided to take the opportunity to use a Python notebook from the University of Ghent's Centre for Digital Humanities.¹⁰⁰ It was created by Bas Verduyck and utilizes Tesseract OCR, which is an open-source optical character recognition (OCR) engine originally developed by HP and now maintained by Google. Like Transkribus, it converts images of text—like scanned documents, photos, or PDFs, etc—into machine-readable text using advanced algorithms to recognize characters and words in multiple languages. While Transkribus offers the ability to fine-tune transcription models and includes numerous built-in options, I opted to use Tesseract for the Python practice and found Tesseract sufficiently accurate when using good quality photos of printed text.

¹⁰⁰ "Team | Ghent Centre for Digital Humanities."

Work	Transcription service
<i>Histoire</i> first edition	Sorbonne U. & manual
<i>Histoire</i> second edition	Sorbonne U. & manual
<i>Histoire</i> third edition	Tesseract OCR
Perrault's <i>Mother Goose Tales</i>	WikiSource
Perrault's <i>La Belle au bois dormant</i>	Sorbonne U. & manual
Choisy's <i>Mémoires pour servir à l'histoire de Louis XIV</i>	Internet Archive
Choisy's <i>Mémoires Habille en Femme</i>	Internet Archive
Choisy's <i>Journal du voyage de Siam</i>	Transkribus OCR
L'Héritier's <i>Œuvres meslées</i>	Transkribus OCR
L'Héritier's <i>Marmoisan</i>	Transkribus OCR

Table 1. Works included with the type of transcription used.

❖ ❖ Creating the Notebook

The Python notebook I created for this project draws upon a variety of resources to support the data processing and analysis tasks involved. I chose Google Colab over Jupyter or Anaconda, two popular local development environments for Python, because Colab provides data integration with Google Drive, which simplifies accessing and storing data. My initial exposure to using Colab came from the Center for Digital Humanities (CDH) at UGhent, where I also learned methods for loading a corpus into the Colab environment.

For data cleaning, including the removal of stop words and punctuation using NLTK in Python, I adapted techniques shared by Geeks for Geeks¹⁰¹, while refining portions of the code with guidance from the Programming Historian¹⁰². For stylometric analysis, specifically calculating Burrows's Delta, I followed a tutorial by Thomas Wood¹⁰³ at Fast Data Science, which offered a clear, practical approach to authorship attribution. Each of these sources

¹⁰¹ "Removing Stop Words with NLTK in Python."

¹⁰² Laramée, "Introduction to Stylometry with Python."

¹⁰³ Wood, "Fast Stylometry."

contributed significantly to the notebook’s development, reflecting the collaborative knowledge shared within the digital humanities and data science communities.

First I installed the library:

```
!pip install faststylometry

faststylometry in /usr/local/lib/python3.10/dist-packages (1.0.4)
numpy (1.24.3)
pandas (2.1.0)
scikit-learn (1.3.0)
wget (3.2)
python-dateutil (2.8.2)
pytz (2024.2)
tzdata (2024.2)
scipy (1.13.1)
joblib (1.4.2)
threadpoolctl (3.5.0)
six (1.16.0)
```

The “pip install” command is a simple way to download and install Python packages from the Python Package Index (PyPI) directly into the user’s development environment. So when I used “pip install <package_name>”, it not only installed the specified package but also checked for any dependencies that package requires, ensuring all necessary components are downloaded and compatible. This helped prevent issues with missing or mismatched dependencies, making it easy to add complex libraries and their requirements to my project with a single command. Then I started downloading the functions that will be needed later:

```
from faststylometry import Corpus
from faststylometry import load_corpus_from_folder
from faststylometry import tokenise_remove_pronouns_en
from faststylometry import calculate_burrows_delta
from faststylometry import predict_proba, calibrate, get_calibration_curve

import nltk
import os
import re
nltk.download("punkt")
from nltk.corpus import stopwords
nltk.download('stopwords')
from nltk.tokenize import word_tokenize
```

In Python, “import” is a keyword used to bring in external modules or libraries into a script, allowing access to pre-written functions, classes, and variables without having to code them from scratch. For example, using “import nltk” gets me access to the Natural Language Toolkit (NLTK), which is one of the largest Python libraries for performing various Natural Language Processing tasks. In my case I also import specific parts of a module (“from nltk.corpus import stopwords”) to access only the features needed, making the code more efficient and modular.

Lastly, I established the notebook’s connection to Google Drive and the files stored there:

```
from google.colab import drive
drive.mount('/content/drive/')
```

And the dataset in Google Drive looks like this:

```
data / train / perrault_-_mothergoose.txt
                perrault_-_labelleauboisdormant.txt
                heritier_-_oeuvresmeslees.txt
                heritier_-_marmoisan.txt
                choisy_-_voyage_de_siam.txt
                choisy_-_memoires.txt
                choisy_-_louisxiv.txt
data / test /  anonymous_-_histoire1ed.txt
                anonymous_-_histoire2ed.txt
                anonymous_-_histoire3ed.txt
```

In Fast Stylometry, the file naming convention is designed to help the code easily identify and separate the author name from the title of each document, which is important for accurate authorship attribution and analysis. Specifically, filenames should use “-” as a separator between the author’s name and the document title. This format ensures consistency in how the files are referenced, parsed, and annotated in the code, making it straightforward to extract the author information and title programmatically. For instance, a file named “data/train/perrault_-_mothergoose.txt” would allow the code to recognize “perrault” as the author and “mothergoose” as the title. This structured approach simplifies the management and

retrieval of metadata across the dataset, supporting tasks like cross-referencing authors and analyzing stylistic differences between works.

✧ ✧ Data Preprocessing

In the data preprocessing stage, I loaded the training texts—those with known authors for comparison. To preserve the original dataset during preprocessing, I renamed it with an “=” symbol, such as “file = train_corpus,” which created a copy of the data. This ensured that any modifications or cleaning steps affected only the duplicate and not the original dataset. This approach safeguards the integrity of the initial dataset, allowing me to reprocess if needed without having to reload the data.

```
train_corpus = load_corpus_from_folder("/faststylometry/data/train")
```

The next step was removing stop words. As a reminder, stop words are a list of common words (e.g., "the," "is," "at") that are typically filtered out to focus on more meaningful vocabulary for authorship attribution. The NLTK library has a variety of languages available in order to select stop words; currently they are Arabic, Danish, Dutch, English, Finnish, French, German, Hungarian, Italian, Norwegian, Portuguese, Romanian, Russian, Spanish and Swedish.

Below are French stopwords available in NLTK:

```
print(stopwords.words('french'))
```

```
['au', 'aux', 'avec', 'ce', 'ces', 'dans', 'de', 'des', 'du', 'elle', 'en', 'et', 'eux', 'il', 'ils', 'je', 'la', 'le', 'les', 'leur', 'lui', 'ma', 'mais', 'me', 'même', 'mes', 'moi', 'mon', 'ne', 'nos', 'notre', 'nous', 'on', 'ou', 'par', 'pas', 'pour', 'qu', 'que', 'qui', 'sa', 'se', 'ses', 'son', 'sur', 'ta', 'te', 'tes', 'toi', 'ton', 'tu', 'un', 'une', 'vos', 'votre', 'vous', 'c', 'd', 'j', 'l', 'à', 'm', 'n', 's', 't', 'y', 'été', 'étée', 'étées', 'étés', 'étant', 'étante', 'étants', 'étantes', 'suis', 'es', 'est', 'sommes', 'êtes', 'sont', 'serai', 'seras', 'sera', 'serons', 'serez', 'seront', 'serais', 'serait', 'serions', 'seriez', 'seraient', 'étais', 'était', 'étions', 'étiez', 'étaient', 'fus', 'fut', 'fûmes', 'fûtes', 'furent', 'sois', 'soit', 'soyons', 'soyez', 'soient', 'fusse', 'fusses', 'fût', 'fussions', 'fussiez', 'fussent', 'ayant', 'ayante', 'ayantes', 'ayants', 'eu', 'eue', 'eues', 'eus', 'ai', 'as', 'avons', 'avez', 'ont', 'aurai', 'auras', 'aura', 'aurons', 'aurez', 'auront', 'aurais', 'aurait', 'aurions', 'auriez', 'auraient', 'avais', 'avait', 'avons', 'aviez', 'avaient', 'eut', 'eûmes', 'eûtes', 'eurent', 'aie', 'aies', 'ait', 'ayons', 'ayez', 'aient', 'eusse', 'eusses', 'eût', 'eussions', 'eussiez', 'eussent']
```

The Fast Stylometry library was designed with only English texts in mind, which was why I needed the addition of the NLTK. You can see the code I used below, and note that any line with a preceding # is meant to be commentary rather than a command.

```
# Setting the language for the stopwords function
stop_words = set(stopwords.words('french'))

# Function to remove stop words from a sentence
def remove_stopwords(sentence):
    words = word_tokenize(sentence)
    filtered_words = [word for word in words if word.lower() not in stop_words]
    return ' '.join(filtered_words)

# Function to load and process each file in the corpus folder
def load_and_clean_corpus(folder_path):
    cleaned_corpus = []
    for filename in os.listdir(folder_path):
        file_path = os.path.join(folder_path, filename)

        # Process only text files (in case the folder has other files)
        if os.path.isfile(file_path) and filename.endswith('.txt'):
            with open(file_path, 'r', encoding='utf-8') as file:
                text = file.read()
                # Remove stop words
                cleaned_text = remove_stopwords(text)
                cleaned_corpus.append(cleaned_text)

    return cleaned_corpus

# Define path to the corpus folder
folder_path = "/content/drive/MyDrive/faststylometry/data/train"

# Load and clean the corpus
cleaned_corpus = load_and_clean_corpus(folder_path)

#Renaming the corpus so it's not needed to redo the code later on
cleaned_corpus = train_corpus
```

The last bit of code above is intended to rename the cleaned training data to a designation that aligns with the naming conventions used in the Fast Stylometry library. By doing this at this

stage, I avoided having to make multiple adjustments across various blocks of code later on, thus streamlining the overall coding process and minimizing potential errors.

The following code snippet is also derived from the Fast Stylometry framework, and it performs a set of tasks similar to the more complicated code block above, but this time for English text. This demonstrates how an optimized function can efficiently handle complex operations based on a single command, showcasing the power of abstraction in programming. Such functions not only simplify the code but also enhance readability and maintainability, allowing for quicker adaptations when working with different languages or datasets.

```
#This is meant for English
train_corpus.tokenise(tokenise_remove_pronouns_en)
```

Next it was time to repeat the same process to the testing data; in this case the three different editions of the Histoire.

```
test_corpus = load_corpus_from_folder("/faststylometry/data/test", pattern=None)

# Setting the language for the stopwords function
stop_words = set(stopwords.words('french'))

# Function to remove stop words from a sentence
def remove_stopwords(sentence):
    words = word_tokenize(sentence)
    filtered_words = [word for word in words if word.lower() not in stop_words]
    return ' '.join(filtered_words)

# Function to load and process each file in the corpus folder
def load_and_clean_corpus(folder_path):
    cleaned_corpus = []
    for filename in os.listdir(folder_path):
        file_path = os.path.join(folder_path, filename)

        # Process only text files
        if os.path.isfile(file_path) and filename.endswith('.txt'):
            with open(file_path, 'r', encoding='utf-8') as file:
                text = file.read()
                # Remove stop words
                cleaned_text = remove_stopwords(text)
                cleaned_corpus.append(cleaned_text)
    return cleaned_corpus
```

```

# Define path to the corpus folder
folder_path = "/content/drive/MyDrive/DATASETS/faststylometry/data/test"

# Load and clean the corpus
cleaned_corpus = load_and_clean_corpus(folder_path)

# Renaming corpus for later code usage
cleaned_corpus = test_corpus

#holding off on this part because it's for English only
test_corpus.tokenize(tokenise_remove_pronouns_en)

```

Like the training corpus, you can see that I also used the tokenize function (visible immediately above as “tokenise_remove_pronouns_en”) even though it was designed for English and not for French. Despite the redundancy of tokenizing twice—once in the primary data cleaning code and then again here—this choice was necessary due to persistent errors that would arise when attempting to continue through the notebook without it. Each attempt to proceed without reapplying tokenization resulted in errors, disrupting the workflow. Therefore, ensuring that the notebook continued running smoothly took precedence, even if it meant tokenizing twice. Ultimately, since the English-specific elements like the pronoun removal in “tokenise_remove_pronouns_en” does not interfere with French-language texts, this workaround did not negatively impact the French corpus in any way.

What tokenizing does is to transform the text into something a machine can understand—strictly speaking, words and letters are meaningless to a computer that only really understands numbers. This transformation of text data into numerical form is referred to as text vectorization, a process that allows machines to treat words and phrases in ways that are computationally meaningful. Tokenization specifically involves breaking down text into manageable, meaningful units—typically individual words, though sometimes phrases or characters depending on the task. Once tokenized, Fast Stylometry employs a “bag of words” model, which disregards grammar and word order, reducing the text to a list of unique words and their frequencies. This reduction to word frequencies not only simplifies the data but also provides the foundation for further analysis.

The bag-of-words model used here is crucial for identifying stylometric features, as it enables the extraction of unique word patterns, authorial markers, and lexical frequencies that

contribute to authorship attribution. Ignoring grammar and word order is advantageous in this context, as the focus is on stylistic elements, such as word choice and frequency, rather than meaning or syntax. By transforming the data in this way, the text becomes ripe for additional machine learning and NLP techniques, which rely on these vectorized representations to establish a structured dataset that computers can readily analyze for patterns and trends.

❖ ❖ **Fast Stylometry for Burrough's Delta**

John Burrows' Delta¹⁰⁴ method is one of the leading stylometric techniques used for authorship attribution today. It calculates the “distance” between a text with unknown authorship and a corpus of texts by known authors, aiming to determine the closest stylistic match. Delta compares the anonymous text to multiple authors’ stylistic signatures simultaneously. Specifically, it measures how both the anonymous text and texts from various known authors deviate from the average of the entire corpus. By assigning equal weight to each feature it measures, the Delta Method mitigates the risk of common words skewing the results, working almost oppositely to visual tools like word clouds. Due to these qualities, Burrows’ Delta method is widely regarded as a robust solution for authorship attribution.

Fast Data Science (FDS) has developed a specialized forensic stylometry model, known as Fast Stylometry¹⁰⁵, that allows users to identify an author based on their distinctive stylistic and linguistic “fingerprint.” This model relies on forensic linguistics principles, where unique patterns in word choice, sentence structure, and syntax create a “fingerprint” that is difficult to replicate. Fast Stylometry is particularly effective when it has access to texts of at least a few thousand words, as larger samples provide a richer dataset for identifying these subtle but consistent stylistic markers.

John Burrows' Delta method is a prominent stylometric technique for authorship attribution, measuring the stylistic "distance" between an anonymous text and a corpus of known authors' works. By comparing deviations from the average style across multiple texts, it minimizes the influence of common words, making it a robust tool for identifying authorship. Fast Data Science has further advanced this field with Fast Stylometry, a model that uses forensic

¹⁰⁴ Laramée, “Introduction to Stylometry with Python.”

¹⁰⁵ Wood, “Fast Stylometry.”

linguistics principles to analyze unique patterns in word choice, sentence structure, and syntax. Fast Stylometry is most effective with larger texts, as they provide a richer dataset for detecting distinctive stylistic markers.

◆◆◆ Vocabulary Range

Before calculating Burrow's Delta and generating graphs and charts, determining the optimal vocabulary size to work with is essential for accurate and meaningful results. Vocabulary size here refers to the number of unique words, or tokens, considered in each text sample. This choice has a significant impact on the analysis, as too small a vocabulary may omit important stylistic markers, while too large a vocabulary can introduce noise, including rare or irrelevant words that do not contribute meaningfully to authorship attribution. By identifying the best vocabulary size, we can ensure that our analysis captures a representative sample of each author's unique stylistic features—such as frequently used words, sentence structures, or lexical choices—while excluding extraneous data that could skew results.

Ultimately, selecting the right vocabulary size enhances the interpretability of the resulting visualizations, such as graphs and dendrograms, making patterns in authorship and stylistic differences clearer. This choice enables us to maintain focus on the core stylistic elements, ensuring that the Burrow's Delta analysis is both precise and capable of capturing the nuances that differentiate one author's work from another.

The optimal vocabulary size for Burrow's Delta varies depending on factors like the text's length, genre, and language. Texts with dense, descriptive language might benefit from a larger vocabulary, while simpler, more repetitive texts might require fewer unique words to capture their essence. Experimentation is often needed to find the best balance, which can involve testing different vocabulary sizes and evaluating their impact on the accuracy and reliability of Burrow's Delta calculations.

In the table below, you can see the experiments I did comparing a known text from each of the three potential authors to the larger training corpus to see what vocabulary size provided the most accurate results. In this specific set of calculations—executed using the function `calculate_burrows_delta(train_corpus, test_corpus, vocab_size=25)`—a lower Delta score indicated a closer match between the texts, thus representing higher accuracy in the results.

Choisy's *Mémoires pour servir à l'histoire de Louis XIV* is represented by "louisxiv," Perrault's *La Belle au bois dormant* by "belle," and Madame L'Héritier's *Marmoisan* or *L'innocente tromperie* by "marmoisan." The result that is best for each author will have the cell(s) highlighted.

Vocab size	louisxiv (Choisy)	belle (Perrault)	marmoisan (Héritier)
25	1.54	1.11	1.26
50	1.36	0.99	0.95
75	1.34	1.02	0.88
100	1.24	1.02	0.78
125	1.18	0.96	0.73
150	1.12	0.94	0.70
175	1.09	1.07	0.78
200	1.08	1.15	0.80
225	1.04	1.12	0.77
250	1.04	1.13	0.78
275	1.00	1.10	0.79
300	0.98	1.11	0.78
325	0.97	1.09	0.80
350	0.95	1.12	0.79
475	0.93	1.32	0.87
575	0.90	1.26	0.89

Table 2. Known Author text compared to that same author's larger corpus.

The results are interesting and highlight that there are similarities in L'Héritier and Perrault's frequency of word usage. The both showed the most accurate results with a vocabulary size of 150 words, making this their ideal size for Burrow's Delta calculations. This means that

for these authors, limiting the analysis to 150 unique words captured the essential stylistic elements needed for reliable attribution, without introducing excessive noise.

Choisy, however, proved to be an outlier. His ideal vocabulary size seems to be significantly larger, at 575 words, suggesting that a broader range of vocabulary better captures the unique stylistic traits in his writing. This difference may indicate that Choisy's style is more complex or varied, requiring more lexical data for accurate analysis. This might also be a reflection of the diverse scope of Choisy's writing included in the corpus, which spans gossiping stories, pious histories, and travelogs. In contrast, L'Héritier and Perrault's contributions to the corpus are more narrowly focused on fairy tales, a genre with its own characteristic simplicity and recurring themes. The broader vocabulary required for Choisy likely reflects his wider thematic range and more complex narrative style, underscoring how genre and content diversity can influence authorship attribution models.

The distinct vocabulary requirements across these authors underscore the importance of tailoring vocabulary size to each writer's unique style, as it enhances the precision of Burrow's Delta in differentiating authors.

◆◆◆ Delta Results

With the optimal vocabulary sizes established, we can now turn to the Delta Scores to observe how effectively these tailored parameters distinguish each author. The following section presents the calculated Delta Scores, illustrating the degree of stylistic similarity between each edition of the *Histoire* and the three potential authors.

```
calculate_burrows_delta(train_corpus, test_corpus, vocab_size = 150)
```

Potential Authors	Histoire 1st Edition	Histoire 2nd Edition	Histoire 3rd Edition
Choisy	1.764114	1.528025	1.507496
Perrault	1.281291	1.133010	1.249633
Héritier	1.319688	1.196863	1.296122

Table 3. Burrows's Delta Method applied to 150 size vocabulary.

Potential Authors	Histoire 1st Edition	Histoire 2nd Edition	Histoire 3rd Edition
Choisy	1.857069	1.629427	1.618432
Perrault	1.394229	1.234385	1.311601
Héritier	1.549789	1.380628	1.435204

Table 4. Burrows's Delta Method applied to 575 size vocabulary.

Although the 575-size vocabulary initially showed promise for Choisy, particularly in the exploratory phase, its performance didn't yield any significant improvements when applied in the Burrows's Delta calculations for the *Histoire* texts. Interestingly, Choisy's Delta scores at this larger vocabulary size did not display a notable decrease, and remained generally consistent across the different editions. This lack of improvement suggests that increasing the vocabulary size beyond 150 words doesn't necessarily enhance the model's accuracy or clarity when distinguishing between potential authors. As such, the larger vocabulary size was primarily included for comparison and exploratory purposes, providing insight into how varying the vocabulary affects attribution accuracy. Ultimately, the comparison highlighted that, while a broader vocabulary can sometimes help reveal nuanced stylistic elements, a carefully chosen smaller vocabulary may be more effective for the specific goal of authorship attribution.

If I were to approach this project again, I would aim to include a broader range of work styles by both Perrault and L'Héritier. Currently, the dataset relies heavily on fairy tales for these two authors, while Choisy's contributions span a much wider thematic and stylistic range—from lighthearted gossip and travelogs to serious historical accounts. This difference in genre variety may skew the results in favor of Perrault and L'Héritier, as the algorithm's focus on stylistic markers may overlook or downplay the stylistic variety present in Choisy's work simply because it doesn't have a comparable range for Perrault or L'Héritier. By expanding the selection to include more varied types of texts from each author, such as Perrault's moral tales or any historical or epistolary writings by L'Héritier, the analysis would provide a more balanced comparison across authors, reducing genre-based biases.

A more varied corpus for all authors would also strengthen the algorithm's ability to recognize broader stylistic patterns that aren't tied exclusively to one genre. This would offer a more accurate and comprehensive view of each author's unique voice and enhance the model's

generalizability across different types of content. Adding this range would not only contribute to more robust authorship attribution but also yield insights into how each author adapted their style across genres, enriching our understanding of their literary contributions.

◆ Findings

The results from Burrow's Delta analysis reveal intriguing patterns in authorship attribution across the three editions of the *Histoire*. Our potential authors are Choisy, Perrault, and L'Héritier, with delta scores calculated for each. Choisy shows delta scores of 1.764114 for the first edition, 1.528025 for the second, and 1.507496 for the third. Perrault, meanwhile, has notably lower scores: 1.281291 for the first edition, 1.133010 for the second, and 1.249633 for the third. L'Héritier's scores are similarly low, with 1.319688 for the first, 1.196863 for the second, and 1.296122 for the third edition. Lower delta scores indicate closer similarity in word usage and vocabulary to the *Histoire* editions, making L'Héritier and Perrault more likely candidates for authorship than Choisy.

Interestingly, the first edition of the *Histoire* consistently yields the highest delta scores for all three authors, suggesting it has fewer stylistic similarities with any of them compared to the later editions. This shift could be due to the fact that the text nearly doubled in length from the first to the second edition, expanding from 7,085 words to 13,374 words. The increase in length may have introduced new vocabulary and stylistic features that align more closely with Perrault and L'Héritier's known writing styles, bringing down the delta scores for the later editions.

While Choisy's scores are generally higher, indicating less similarity, the concern was that his results could be skewed due to the diversity in his writings, which include memoirs, historical accounts, and travel narratives rather than fairy tales. However, if Choisy's genre truly influenced the results unfavorably, we might expect Perrault and L'Héritier's scores to be nearly identical given their similar genre and familial connection. Yet, the delta scores for Perrault and L'Héritier differ slightly across the editions, implying that their styles are not as interchangeable as one might assume. This distinction provides a clearer case for why both Perrault and L'Héritier may share stylistic links with the *Histoire*—a factor that further strengthens the hypothesis of shared authorship or influence within their literary circle.

Taking another look at the results, this time presenting them in chart format for clearer, more immediate interpretation. While visualizing the data, we can more easily observe trends and differences in delta scores across authors and editions, which may reveal additional insights that are less obvious in the numerical format alone

```

from matplotlib import pyplot as plt

# Define the figure and its size
plt.figure(figsize=(8, 4))

# Plot each series with a label for identification in the legend
_df_7['anonymous - histoire1ed'].plot(kind='line', label='anonymous - histoire1ed')
_df_7['anonymous - histoire2ed'].plot(kind='line', label='anonymous - histoire2ed')
_df_7['anonymous - histoire3ed'].plot(kind='line', label='anonymous - histoire3ed')

# Customize the plot
plt.title('Comparison of Text Examples')
plt.gca().spines[['top', 'right']].set_visible(False)
plt.legend() # Adds a legend to identify each line

# Show the plot
plt.show()

```

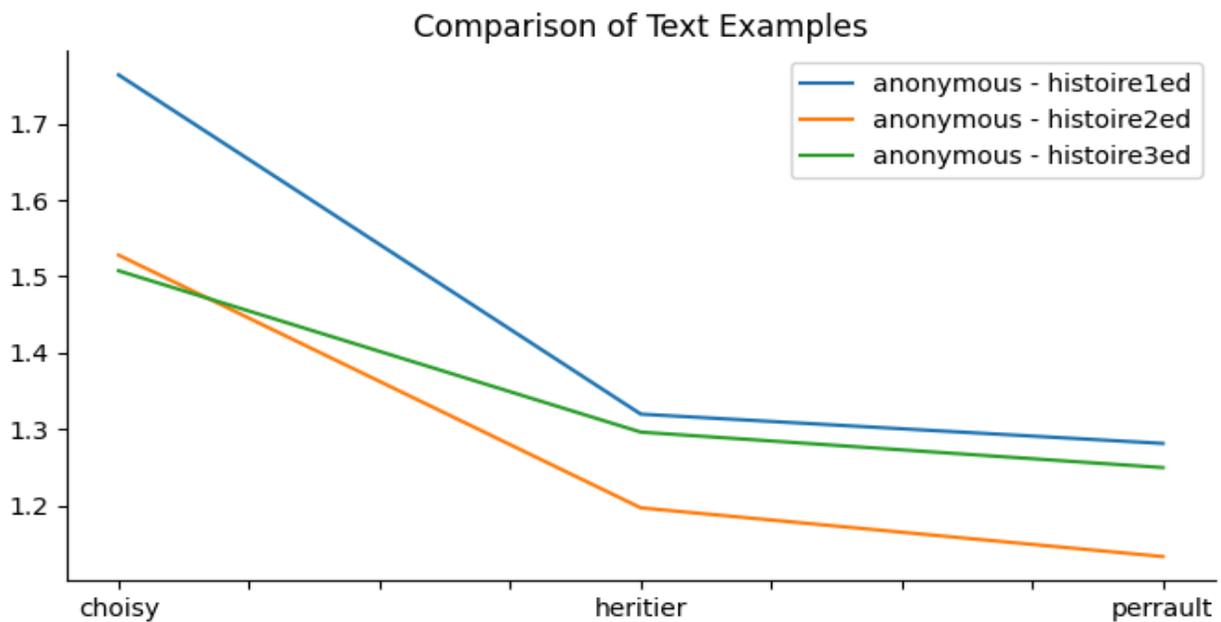


Figure 13. Line chart created in Python.

In reviewing the chart, it becomes strikingly clear that Perrault's writing style aligns most closely with the second and third editions of the *Histoire de la Marquise-Marquis de Banneville*. The modifications introduced in the third edition—such as added scenes and other content adjustments—seem to mirror Perrault's stylistic tendencies even more closely. This suggests a potential editorial influence or at least a strong stylistic similarity, especially in the extended sections. However, the results are not as definitive as one might hope for in authorship attribution; the data leans toward Perrault and L'Héritier but does not entirely rule out other contributions or influences.

Interestingly, while the findings imply that Choisy likely wasn't directly involved in the writing, this does not exclude the possibility of his influence through conceptual contributions, discussions, or even advisory input on the narrative. Choisy's known association with the *Histoire* in certain historical circles might still indicate his influence on the text's thematic or narrative elements, even if he was not the principal author. This opens an intriguing avenue for considering collaborative or dialogical influences on the work, which, while less visible through Burrow's Delta, might yet shape how the text was developed in its final form.

Conclusion

Queer narratives and identities are often viewed as recent constructs, but centuries-old myths and fairy tales reveal enduring themes of fluid gender, transformation, and self-discovery, connecting modern queerness to a deep-rooted storytelling tradition. Mythology, with its gods of shifting identities and transformative powers, flowed naturally into folk tales that retained these themes within human-centered narratives, reflecting societal values through familiar characters like tricksters and gender-fluid beings. As fairy tales evolved into a literary genre, they continued to address questions of identity, transformation, and belonging, often in subversive ways despite their moral closure and appeal to elite audiences. This connection between fairy tales and queerness, explored through the *Histoire de la Marquise-Marquis de Banneville*, inspired my research in Digital Humanities.

Leveraging tools like stylometry to analyze authorship, I sought to investigate whether computational techniques could resolve the question of the *Histoire*'s anonymous authorship, believed by some scholars to be linked to Perrault, L'Héritier, or Choisy. By applying programming skills and digital analysis to this obscure, distinctly queer tale, I aimed to bridge classic literature and modern technology, enriching both the understanding of the *Histoire* and the digital methodologies in literary studies.

When I first encountered the *Histoire de la Marquise-Marquis de Banneville*, it immediately resonated with me as a thesis topic for my Master's in Rare Books and Digital Humanities. The story's origins in *Mercure Galant*, a 17th-century Parisian periodical influential among elite readers, particularly educated women, hinted at a unique social and literary context that engaged with gender and class norms. The *Mercure Galant*, founded by Donneau de Visé, showcased a diverse range of content, including poetry and social critiques, making it a cultural touchstone of late seventeenth-century France. The *Histoire de la Marquise-Marquis de Banneville* gained popularity, leading to reprints that reflected its significance. Inspired by this context, I adopted an interdisciplinary approach, combining rare book history with digital analysis to explore the story's history and scholarly reception. Through my research, I aimed to deepen the understanding of how stories with progressive themes circulated in early modern Europe, suggesting a more nuanced view of identity and narrative diversity than is often assumed.

The authorship of the *Histoire de la Marquise-Marquis de Banneville* has been attributed to three different individuals over the past three centuries, with significant scholarly discourse emerging primarily in the last century. In 1906, Paul Bonnefon proposed that the tale was likely written by Madame L'Héritier, a relative of Charles Perrault, rather than Perrault himself. Later, in 1928, Madame Jeanne Roche-Mazon suggested that the story might be a collaborative work between Perrault and the abbé de Choisy, adding complexity to the authorship debate. In 2004, the Modern Language Association published a standalone edition of the *Histoire*, which included an introduction by Joan DeJean that examined the potential authorship by considering various candidates, including the possibility of collaboration.

Despite the scholarly discussions, no definitive consensus has emerged, and notably, previous analyses have not utilized statistical or computational methods to explore these claims. This presents an opportunity for a new approach to authorship studies through modern text

analysis techniques, aiming to uncover patterns that could clarify authorship and illuminate the broader context of gender and collaboration in late seventeenth-century literature. Ultimately, this research seeks to enhance the understanding of the *Histoire*'s significance within the fairy tale canon and its place in literary queer history.

It's true that determining authorship can resemble a literary mystery, requiring careful detective work to uncover the individual behind the words. Historically, authorship was less of a concern, with many ancient texts—like clay tablets and Greek philosophical scrolls—being circulated anonymously, as the focus remained on the content rather than the creator. It is only in recent centuries that authorship has gained importance, particularly with the rise of intellectual property and the celebration of the individual author. Today, while clear attribution is often expected, many works remain anonymous or pseudonymous, presenting challenges for scholars and historians. Establishing authorship typically relies on evidence and analysis; clear external evidence, like signatures, can confirm authorship, while contextual clues can help deduce the author in their absence.

In more ambiguous cases, quantitative analysis and stylometry—the statistical study of writing style—have become essential tools. This modern approach allows researchers to analyze texts for linguistic features such as word frequency and sentence structure, enabling comparisons to identify potential authors. By preparing and structuring data for stylometric analysis, I hoped to uncover patterns that could reveal the true authorship of the *Histoire de la Marquise-Marquis de Banneville* by examining texts from its three potential authors: de Choisy, Perrault, and L'Héritier.

My process of authorship attribution began with the selection and acquisition of texts related to the *Histoire de la Marquise-Marquis de Banneville*, focusing on early versions available at the BnF. These included both digital copies of first and second editions and a print version of the 1723 edition. After gathering texts from the three potential authors, my next step involved digitizing and preprocessing the data to ensure it was clean and uniform. This preprocessing entailed removing special characters and correcting any errors from OCR. The data was then structured using Python. The aim was to compare the linguistic styles of the *Histoire* with contemporaneous works by the three potential authors, ensuring a fair comparison based on themes, style, and historical context. Digital versions were sourced through the BnF's

Gallica website the Internet Archive, and Project Gutenberg, while physical editions were captured using a scan tent for comprehensive analysis.

For the works that were not yet transcribed or accessible online, I used PDFs converted to PNG images and personal digital photos to create machine-readable texts via Transkribus and Tesseract OCR. Transkribus, known for its advanced transcription capabilities and user-friendly interface, was initially employed to transcribe L'Héritier's *Œuvres meslées* and Choisy's *Journal du voyage de Siam*. As I progressed in my studies, they gained familiarity with Python and utilized Tesseract OCR for processing the *Histoire's* third edition images, finding it accurate for good-quality printed text.

I created a Python notebook for data processing and analysis using Google Colab for its integration with Google Drive, which simplifies data access and storage. I adapted data cleaning techniques from various online resources, employing NLTK for tasks like stop word removal and using the Fast Stylometry library for stylometric analysis, specifically Burrows's Delta. The preprocessing involved removing common stop words and tokenizing the text, despite some redundancy, to maintain a smooth workflow. This text transformation, particularly through tokenization and the bag-of-words model, allows for effective analysis of stylistic features and authorial markers, setting the stage for further machine learning and natural language processing applications.

With the optimal vocabulary size established, my analysis shifted to the Delta scores, which assess how well these tailored parameters differentiate among the authors of the *Histoire*. The calculated Delta Scores reveal the degree of stylistic similarity between each edition of the *Histoire* and the potential authors, Choisy, Perrault, and Hérítier, when applying a 150-word vocabulary size. The results indicate that increasing the vocabulary size beyond this does not enhance authorship attribution accuracy, suggesting that a carefully selected smaller vocabulary may be more effective for this specific task.

The results of Burrow's Delta analysis reveal significant insights into authorship attribution for the three editions of the *Histoire*, focusing on potential authors Choisy, Perrault, and L'Hérítier. Delta scores indicate that Perrault and L'Hérítier exhibit notably closer stylistic alignment with the *Histoire* editions compared to Choisy. The first edition consistently shows the highest delta scores across all authors, implying fewer stylistic similarities, likely due to its shorter length compared to the later editions. This increase in length may have introduced new

vocabulary and features that align more closely with the styles of Perrault and L'Héritier. Although Choisy's higher scores suggest less similarity, they may be influenced by the diverse genres of his works, which include memoirs and travel narratives.

Interestingly, while Perrault and L'Héritier share similarities in genre, their differing delta scores indicate distinct stylistic traits. A line chart visualization further supports the findings, demonstrating that Perrault's writing style closely aligns with the second and third editions, suggesting a possible editorial influence. Although the results lean toward Perrault and L'Héritier, they do not completely exclude the possibility of Choisy's influence, hinting at potential collaborative contributions to the text's thematic elements.

In conclusion, my study of the *Histoire de la Marquise-Marquis de Banneville* combines computational methods with literary analysis to explore authorship in the context of early modern queer narratives. By examining delta scores from stylometric analysis, I found that Perrault and L'Héritier's styles aligned more closely with the *Histoire* than Choisy's, suggesting that either or both authors may have contributed to the text. Interestingly, the analysis reveals that the first edition is the least stylistically similar to all three potential authors, perhaps due to its shorter length, which lacks the expanded narrative features present in later editions. This suggests that the text may have undergone editorial modifications that reflected the stylistic preferences of its possible authors over time, particularly in the second and third editions.

This computational analysis method could serve as a powerful tool for scholars of queer texts, encouraging nuanced discussions around authorship and identity in historical literature. By making my code openly available on my website, I aim to enable other researchers to apply these techniques to a range of texts, opening new possibilities for examining themes of gender and transformation across literary history. I hope this will inspire a cascading effect, where each study contributes to a growing body of scholarship that reexamines classic works with a fresh lens. I'll be making the Python notebooks available to humanities scholars on my website, presented in the Digital Object chapter after this one.

Future research could delve deeper into genre-specific influences and expand the analysis to include a broader corpus of texts from other contemporaries of Perrault, L'Héritier, and Choisy, which might yield further insight into stylistic trends within this literary circle. Additionally, applying machine learning techniques to authorship attribution could offer new perspectives on subtle stylistic markers, enriching our understanding of anonymous and

collaborative authorship in early literary works. Ultimately, this research illuminates the intricate relationship between narrative form and authorial influences in historical texts, showing the potential for digital humanities to enhance both literary scholarship and our understanding of early queer identities.

Digital Object

I chose to make a website from scratch rather than using Wordpress,¹⁰⁶ I have experience in using HTML for text encoding and I've enjoyed expanding on my CSS skills during the course of this masters program. HTML and CSS are the foundational languages for creating websites, each playing a distinct role in how a site looks and functions.

HTML (HyperText Markup Language) is a markup language that structures the content of a webpage. Think of it as the backbone of a website: it organizes text, images, links, and other elements into a coherent layout by using a series of tags and elements. For instance, headings, paragraphs, and lists in HTML give the webpage its structure and organization, while embedding the meaning of each part so web browsers can interpret and display it correctly. CSS (Cascading Style Sheets), on the other hand, is used to style and visually enhance these HTML elements. CSS allows control over colors, fonts, layout, spacing, and nearly every visual aspect of a webpage, transforming a basic HTML structure into a polished, cohesive, and user-friendly interface. It also enables designs to adapt across different screen sizes and devices.

Choosing to code directly in HTML and CSS offers more granular control than a WordPress site, where many design aspects are handled through themes and plugins. Although WordPress is highly popular for its ease of use and extensive features—especially for users needing quick setups and pre-designed layouts—coding from scratch provides complete creative

¹⁰⁶ This choice may be The *Histoire de la Marquise-Marquis de Banneville* is one such story. Written in the late Seventeenth-century, it tells a fictional tale about a mother who is willing to do anything to keep a son from ever going to war, and a daughter who is more than what she seems. It has a surface-level similarity to other familiar fairy tales like *Cinderella*; a girl from the countryside being introduced to the court, a family friend playing the role of a fairy godmother to buy her dresses and take her to the balls, and a handsome love interest that sweeps our heroine off her feet. However, we have something wildly unique about the *Histoire*, and that is the story's repeated motif of gender presentation being separate from a character's assigned sex. prescient considering the difficulties Wordpress is currently facing with their integrated plugins.

and functional flexibility, allowing precise customization without the constraints of pre-made themes or third-party dependencies. This direct approach is particularly beneficial for learning web development, as it enhances familiarity with core web concepts, and it's especially useful for complex projects where unique styling or functionality is required.

For guidance on best practices and CSS tricks, I found W3Schools a helpful resource. It provides clear explanations and examples for HTML, CSS, and other web technologies, making it useful for both beginners and those looking to refine specific techniques. There are also a few elements of Javascript that I incorporated. JavaScript is a programming language used to add interactivity and dynamic behavior to websites. While HTML provides the structure and CSS handles the styling, JavaScript brings websites to life by allowing users to interact with elements on the page. For example, JavaScript can make buttons clickable, validate forms, update content dynamically without reloading the page, and create animations.

◆ Design & Layout

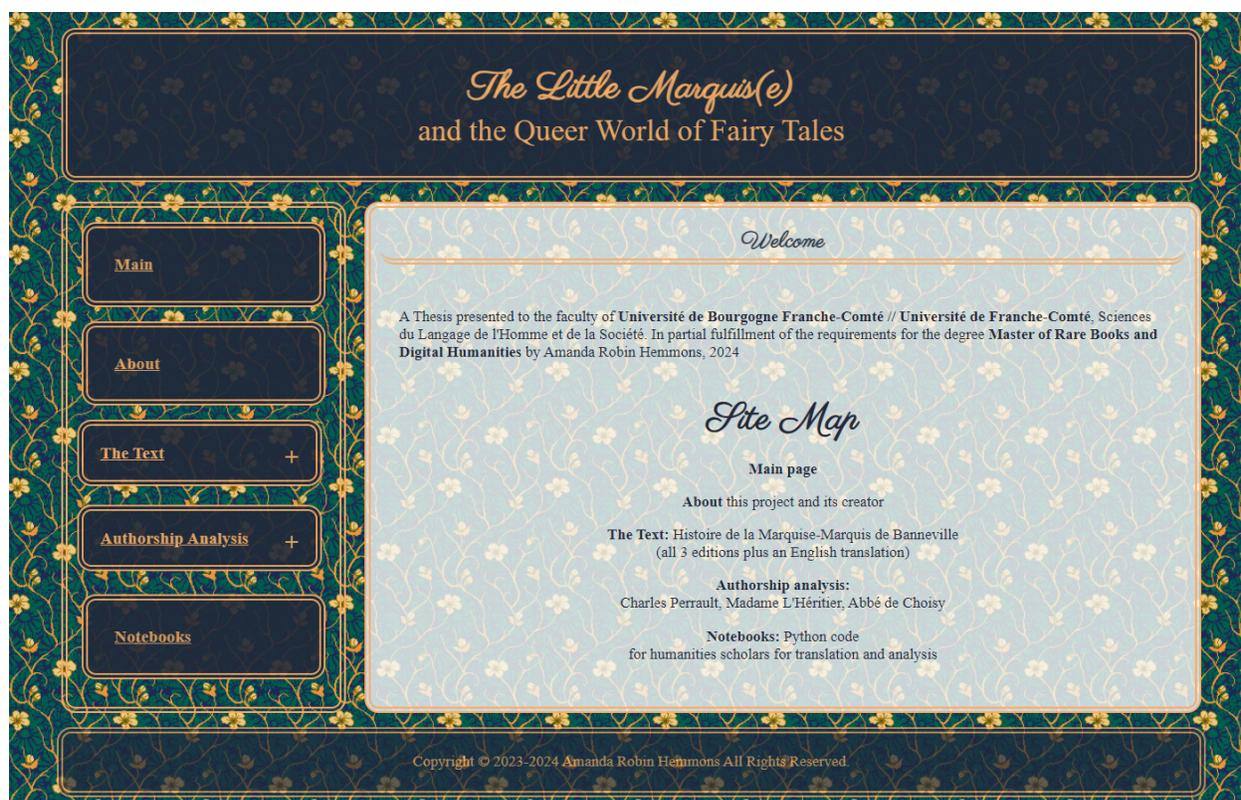


Figure 14. Digital object/Website, located at <https://inkshield.org/histoirethesis/index.html>

The background is a repeating image I sourced from a Neocities site that curates retro-inspired web design elements. My goal was to echo the Baroque period's intricate filigree and ornamental details, reflecting the aesthetic sensibilities popular when *Histoire de la Marquise-Marquis de Banneville* was published. Baroque design emphasizes elaborate patterns, rich textures, and layered details, all of which present a unique challenge to recreate digitally. A true Baroque look would involve ornate scrollwork and gilded accents, features that are difficult to fully capture online without high-resolution images or complex web design frameworks.

To approximate this look within digital constraints, I chose a floral background as a subtle nod to Baroque motifs. I also integrated the Parisienne font in select areas to bring in a sense of elegance reminiscent of the period's typography. Additionally, I applied a colored border radius around various text blocks to mimic the appearance of a gold frame, evoking the era's luxurious textures and framed artwork. Each of these choices helps to recreate an immersive, period-inspired atmosphere, balancing historical authenticity with functional web design.

On the vertical navigation bar along the left side, I incorporated two JavaScript-enabled buttons that expand the menu when clicked. This interactive button allows the navigation bar to stay compact and streamlined by default, but when users click the button, the full menu opens, providing access to all options without taking up too much screen space. This feature enhances usability, particularly on smaller screens or for users who prefer a less cluttered layout.

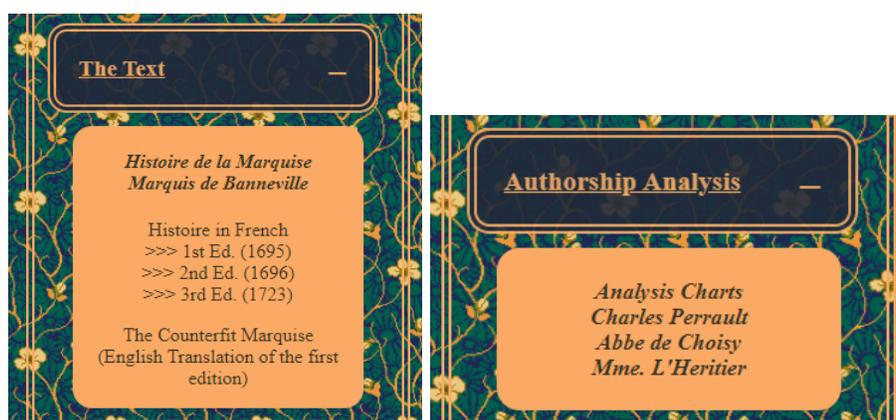


Figure 15. Side navigation drop-down¹⁰⁷ menus.

¹⁰⁷ “How To Create a Side Navigation Dropdown.”

In the footer, I used a JavaScript snippet to automatically display the current year within the Copyright statement, updating it in real time. This means that every January, the year refreshes without any manual changes, ensuring that the site always looks up-to-date. This small but helpful script saves maintenance time, adds a polished touch, and reflects the thoughtfulness that can be achieved with even simple JavaScript functions. Together, these elements make the site feel more dynamic and user-friendly by combining functionality with aesthetic considerations.



Figure 16. Javascript-enabled¹⁰⁸ copyright statement.

◆ Content

Along with the main page, which features a site map, I included four core sections: *About*, *The Text*, *Authorship Analysis*, and *Notebooks*.

The *About* page provides a brief overview of the project, background on the text, and the inspiration behind the site. In *The Text* section, I compiled transcriptions of all three editions of the *Histoire de la Marquise-Marquis de Banneville* for easy comparison and research. The first and second editions were transcriptions begun by Sorbonne University’s *Mercure Galant* project, which I finalized by cross-referencing with the original digitized books. The third edition was the result of OCR processing with Tesseract, which I carefully edited to ensure accuracy.

On the *Authorship Analysis* page, I added biographies of the three potential authors and displayed the results of my computational analysis. This page walks through the authorship debate with insights from my work using stylometry and data analysis, presenting findings in both text and visual form to enhance understanding of the evidence behind each authorship possibility.

¹⁰⁸ “How to Automatically Update Current Year with JavaScript.”

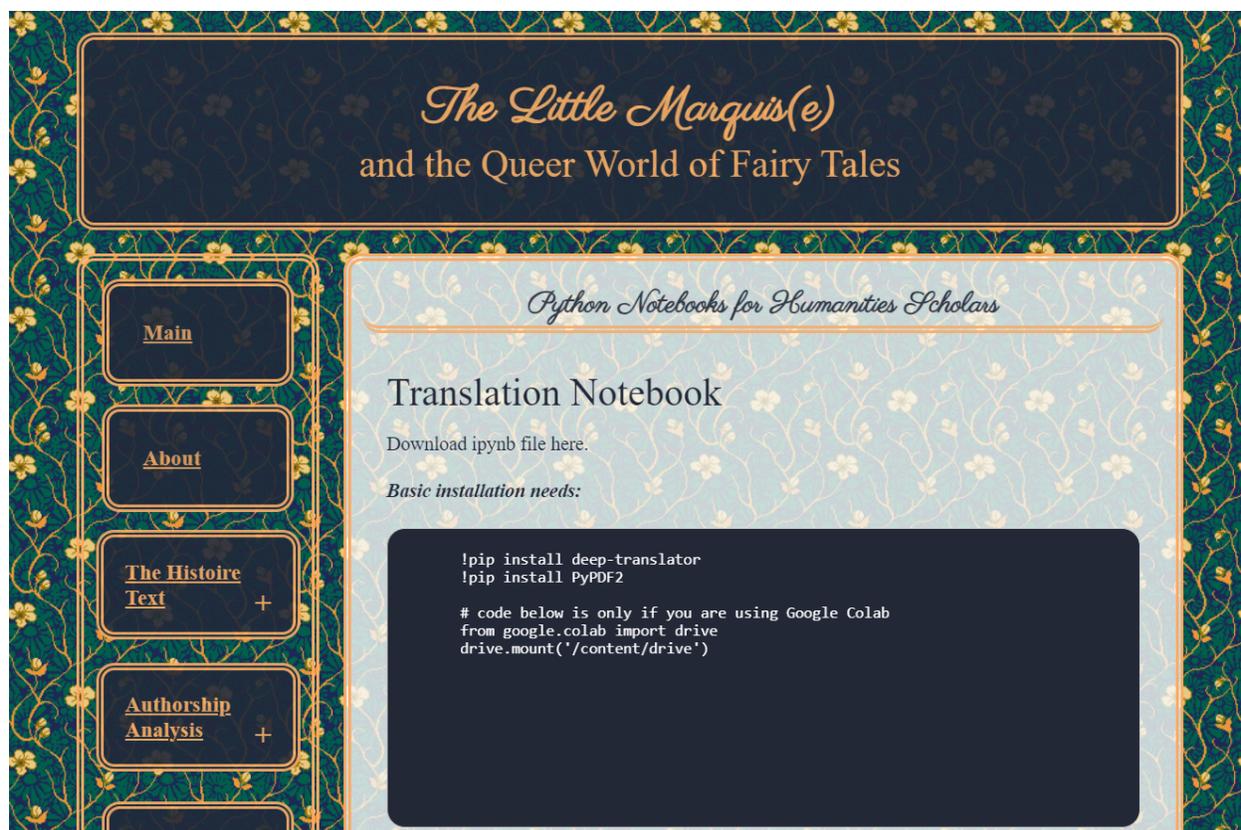


Figure 17. Screenshot of Notebook page.

The Notebooks section is designed as a resource for digital humanities scholars interested in text analysis. Here, I've shared my Python notebooks and useful code snippets. I've included my notebook for translating texts into English and the main analysis notebook. The main analysis notebook loads the *Histoire*, or a scholar's chosen data, as well as the comparison texts to run Burrow's Delta using the Fast Stylometry library.

Each complete Notebook file is available for download, and key sections of code are displayed in scrolling text boxes for easy copying and pasting. Through these practical examples, I aim to encourage others to use Python in their own research and explore new possibilities in text-based literary analysis projects.

Bibliography

1. Adem. *Photo: Japan Themed Editathon in Istanbul Bilgi University Wikipedia Student Club I*. November 30, 2023. Own work.
https://commons.wikimedia.org/wiki/File:Japan_themed_editathon_in_Istanbul_Bilgi_University_Wikipedia_Student_Club_I_12.jpg.
2. Andersen, H. C. (Hans Christian). “Hans Andersen’s Fairy Tales. Second Series.”
<https://www.gutenberg.org/files/32572/32572-h/32572-h.htm>. Accessed October 26, 2024. <https://www.gutenberg.org/cache/epub/32572/pg32572-images.html>.
3. Aulnoy, Marie-Catherine Le Jumel de Barneville baronne d’ (1650-1705) Auteur du texte. *Contes Des Fées, Par Mme d’Aulnoy, Revue Par Mlle Marie Guerrier de Haupt...*, 1868. <https://gallica.bnf.fr/ark:/12148/bpt6k6581768x>.
4. author, Unknown authorUnknown. *L’abbé de Choisy En Femme*. 1855.
http://next.liberation.fr/culture/2014/08/11/la-folle-vie-de-l-abbe_1079103.
https://commons.wikimedia.org/wiki/File:Abb%C3%A9_de_Choisy.jpg.
5. ———. *L’abbé de Choisy En Femme. Illustration*. December 15, 2016.
http://next.liberation.fr/culture/2014/08/11/la-folle-vie-de-l-abbe_1079103.
https://commons.wikimedia.org/wiki/File:Abb%C3%A9_de_Choisy.jpg.
6. Besson, Anne. “From the Aristocracy to the Playground: A Short History of Fairy Tales | Fantasy - BnF.” Accessed October 27, 2024.
<https://fantasy.bnf.fr/en/understand/aristocracy-playground-short-history-fairy-tales>.
7. “BnF - Portrait de Mademoiselle L’Héritier.” Accessed October 29, 2024.
https://classes.bnf.fr/essentiels/grand/ess_1008.htm.
8. Boccaccio, Giovanni. *The Decameron*. The Villon Society, London, 1886.
<https://www.gutenberg.org/files/23700/23700-h/23700-h.htm>.
9. Bonnefon, Paul. “Les Dernières Années de Charles Perrault.” *Revue d’Histoire Littéraire de La France* 13, no. 4 (1906): 606–57.
10. Bottigheimer, Ruth B. *Fairy Tales: A New History*. Albany, N.Y.: Excelsior Editions/State University of New York Press, 2009. <http://site.ebrary.com/id/10588786>.
11. Canva. “Canva.” Accessed October 27, 2024. <https://www.canva.com/>.

12. Carpenter, Humphrey, and Mari Prichard. *The Oxford Companion to Children's Literature*. Oxford University Press, 1984.
<https://archive.org/details/oxfordcompanion00hump/>.
13. C.H., Swaroop. "About Python · A Byte of Python." Accessed October 27, 2024.
https://python.swaroopch.com/about_python.html.
14. "Charles Perrault | Encyclopedia.Com." Accessed October 27, 2024.
<https://www.encyclopedia.com/children/academic-and-educational-journals/charles-perrault>.
15. Choisy, abbé de. "Aventures de l'abbé de Choisy habillé en femme."
<https://www.gutenberg.org/files/54035/54035-h/54035-h.htm>. Accessed October 27, 2024. <https://www.gutenberg.org/cache/epub/54035/pg54035-images.html>.
16. ———. *Mémoires pour servir à l'histoire de Louis XIV*. A Utrecht : Chez Wan-De-Water, 1727. <http://archive.org/details/mmoirespourser00choi>.
17. Choisy, François-Timoléon de. *Histoire de Madame La Comtesse Des Barres*, 1807.
<https://gallica.bnf.fr/ark:/12148/bd6t5751462z/>.
18. ———. *Mémoires de l'abbé de Choisy habillé en femme ; suivi de Histoire de la marquise-marquis de Banneville*. Petite bibliothèque Ombres. Toulouse: Ombres, 1995.
19. Choisy, François-Timoléon de (1644-1724) Auteur du texte. *Aventures de l'abbé de Choisy Habillé En Femme / Précédé d'une Notice et de Documents Inédits Par Marc de Montifaud ; et d'une Eau-Forte de Hanriot*, 1884.
<https://gallica.bnf.fr/ark:/12148/bpt6k2079262>.
20. ———. *Histoire de l'Église. Tome 1 / , Par M. l'abbé de Choisy,...*, 1703.
<https://gallica.bnf.fr/ark:/12148/bpt6k54250306>.
21. ———. *Journal Du Voyage de Siam Fait En M. DC. LXXXV. et M. DC. LXXXVI. Par M. L. D. C.*, 1687. <https://gallica.bnf.fr/ark:/12148/bpt6k8727041z>.
22. Conner, Randy P., David Hatfield Sparks, Mariya Sparks, and Gloria Anzaldúa. *Cassell's Encyclopedia of Queer Myth, Symbol, and Spirit: Gay, Lesbian, Bisexual, and Transgender Lore*. London ; Herndon, VA: Cassell, 1997.
23. Dakota, P. J. *Photo: Westminster Girls' Hockey Players Smile for the Camera during a Game at Jackson Rink*. January 25, 2015. Photo taken during a Girls Hockey game at

- Westminster School's Jackson Rink (Connecticut USA).
https://commons.wikimedia.org/wiki/File:Westminster_Girls_Hockey_players.jpg.
24. De Choisy, Abbie. *The Transvestite Memoirs and the Story of the Marquise-Marquis De Banneville*. London: Peter Owen, 1973.
 25. Delarue, Paul. *Le conte populaire français: contes merveilleux : supplément au Catalogue de Paul Delarue et Marie-Louise Tenèze*. Edited by Bénédicte Bonnemason. Amphi 7. Toulouse: Presses universitaires du Midi, 2017.
http://bvbr.bib-bvb.de:8991/F?func=service&doc_library=BVB01&local_base=BVB01&doc_number=030023095&sequence=000001&line_number=0001&func_code=DB_REC_ORDS&service_type=MEDIA.
 26. Dietrich, Felix. "OCR vs. HTR or 'What Is AI, Actually?'" READ-COOP. Accessed October 27, 2024. <https://readcoop.eu/insights/ocr-vs-htr/>.
 27. Fleming, Patrick C. "Hans Christian Andersen's Sexuality." *Patrick C. Fleming* (blog), August 28, 2018. https://www.pcfleming.com/2018/08/28/andersen_sexuality/.
 28. GeeksforGeeks. "Introduction to NLTK: Tokenization, Stemming, Lemmatization, POS Tagging," December 4, 2022.
<https://www.geeksforgeeks.org/introduction-to-nltk-tokenization-stemming-lemmatization-pos-tagging/>.
 29. GeeksforGeeks. "Removing Stop Words with NLTK in Python," May 22, 2017.
<https://www.geeksforgeeks.org/removing-stop-words-nltk-python/>.
 30. Google Fonts. "Parisienne." Accessed October 30, 2024.
<https://fonts.google.com/specimen/Parisienne>.
 31. Greenberg, Jane. "The Applicability of Natural Language Processing (NLP) to Archival Properties and Objectives." *The American Archivist* 61, no. 2 (1998): 400–425.
 32. Greetham, David. "'Chapter 25: The Resistance to Digital Humanities | David Greetham' in 'Debates in the Digital Humanities' | Debates in the Digital Humanities." Accessed October 27, 2024.
<https://dhdebates.gc.cuny.edu/read/untitled-88c11800-9446-469b-a3be-3fdb36bfd1e/section/4110ccc9-920a-47a2-80eb-77a43f3d6e86>.
 33. Grimm, Wilhelm, and Jacob Grimm. *Kinder- und Haus-Märchen*. 2e Aufl., 1819.
<http://nrs.harvard.edu/urn-3:FHCL.HOUGH:11190297?buttons=y>.

34. “Grimm’s Fairy Tales | History & Best-Known Works | Britannica,” September 24, 2024.
<https://www.britannica.com/topic/Grimms-Fairy-Tales>.
35. “Grimm’s Household Tales, Volume 1 - Wikisource, the Free Online Library,” 1884.
https://en.wikisource.org/wiki/Grimm%27s_Household_Tales,_Volume_1.
36. *Histoire de la marquise, marquis de Banneville*. 3rd ed. Paris: d’Houry, 1723.
37. “How to Automatically Update Current Year with JavaScript.” Accessed October 30, 2024.
<https://www.techstacker.com/how-to-automatically-update-current-year-with-javascript>.
38. “How To Create a Side Navigation Dropdown.” Accessed October 30, 2024.
https://www.w3schools.com/howto/howto_js_dropdown_sidenav.asp.
39. Internet Archive. “Journal Du Voyage de Siam Fait En 1685. & 1686 : Choisy, Abbé de, 1644-1724 : Free Download, Borrow, and Streaming.” Accessed October 28, 2024.
https://archive.org/details/ChoisyJourVoySiam1685goog00_201402.
40. Isnaaa. *Photo: A Picture of Üsküdar University Wikipedia Student Club’s Founder Chairman*. March 21, 2023. Own work.
https://commons.wikimedia.org/wiki/File:Wikipedia_Student_club%E2%80%99s_Chairman.jpg.
41. Jayaswal, Vaibhav. “Text Vectorization: Bag of Words (BoW).” Medium, September 28, 2020.
<https://towardsdatascience.com/text-vectorization-bag-of-words-bow-441d1bfce897>.
42. Kahan, Jeffrey. “‘I Tell You What Mine Author Says’: A Brief History of Stylometrics.” *ELH* 82, no. 3 (2015): 815–44.
43. Keerio, Umed Ali. “Why Deep Learning Is so Called Data-Hungry.” Accessed October 27, 2024.
<https://www.linkedin.com/pulse/why-deep-learning-so-called-data-hungry-umed-ali-keerio>.
44. Lang, Andrew, Henry Justice Ford, and George Percy Jacomb Hood. *The Blue Fairy Book. Edited [with an Introduction] by Andrew Lang. With Numerous Illustrations by H.J. Ford and G.P. Jacomb Hood. L.P. Longmans & Co.: London, 1889.*

45. Laramée, François Dominic. “Introduction to Stylometry with Python.” *Programming Historian*, April 21, 2018.
<https://programminghistorian.org/en/lessons/introduction-to-stylometry-with-python>.
46. laurasimonsdaughter. “The Counterfeit Marquise A Literary Fairy Tale...” *Time for a Story?* (blog). Accessed October 26, 2024.
<https://laurasimonsdaughter.tumblr.com/post/655904037059903489/the-counterfeit-marquise-a-literary-fairy-tale>.
47. Laurea, Tesina di. “Loki’s Gender Fluidity and Bisexuality in Norse Mythology and in Modern Adaptations,” n.d.
48. L’Héritier de Villandon, Marie-Jeanne (1664-1734) Auteur du texte. *Oeuvres Meslées, Contenant L’innocente Tromperie, L’avare Puni, Les Enchantemens de l’éloquence, Les Aventures de Finette, Nouvelles , et Autres Ouvrages, En Vers et En Prose, de Mlle L’H*** [L’Héritier de Villandon], Avec Le Triomphe de Mme Des Houlières, Tel Qu’il a Été Composé Par Mlle L’H****, 1696. <https://gallica.bnf.fr/ark:/12148/bpt6k62376r>.
49. Lys, Charlotte Trinquet du. “Women Soldiers’ Tales during Louis XIV’s War Conflicts.” *Marvels & Tales* 33, no. 1 (2019): 140–56.
<https://doi.org/10.13110/marvelstales.33.1.0140>.
50. MacDonald, George. *The Princess and the Goblin. With Numerous Illustrations*. Philadelphia: J.B. Lippincott & Co., 1872.
51. Magazine, Smithsonian, and Sonja Anderson. “Researchers Decipher Cuneiform Tablet—and Discover It’s a Furniture Receipt.” *Smithsonian Magazine*. Accessed October 27, 2024.
<https://www.smithsonianmag.com/smart-news/researchers-decipher-cuneiform-tablet-and-discover-its-a-furniture-receipt-180984805/>.
52. Malone, Edmond. *A Dissertation on the Three Parts of King Henry VI. Tending to Shew That Those Plays Were Not Written Originally by Shakspeare: By Edmond Malone, Esq.* from the Press of Henry Baldwin, 1787.
53. Marina Warner, ed. *Wonder Tales : Six Stories of Enchantment*. London : Vintage, 1996.
<http://archive.org/details/wondertalessixst0000unse>.
54. Master Rare Book and Digital Humanities. “Master Rare Book and Digital Humanities.” Accessed October 26, 2024. <https://rare-book-and-digital-humanities.ubfc.fr/?lang=en>.

55. *Mercuré galant*. “Histoire de la Marquise Marquis de Banneville.” February 1, 1695. Gallica.
56. *Mercuré galant*. “Histoire de la Marquise Marquis de Banneville.” August 1, 1696. Gallica.
57. *Mercuré galant*. “Histoire de la Marquise Marquis de Banneville.” September 1, 1696. Gallica.
58. *Mercuré galant*. “Histoire de la Marquise Marquis de Banneville.” February 1, 1928. Gallica.
59. “Mercuré Galant | OBVIL.” Accessed October 27, 2024. <http://132.227.201.10:8086/projets/mercure-galant>.
60. Metz Howard, Christine. “Professor Casts Doubt on One of History’s Greatest Cross-Dressing Memoirs.” The University of Kansas, January 2, 2015. <https://news.ku.edu/2015/01/02/ku-professor-casts-doubt-one-historys-greatest-cross-dressing-memoirs>.
61. Miller, Minnie M. “The French Periodical Press during the Reign of Louis XIV.” *The French Review* 5, no. 4 (1932): 301–8.
62. Moureau, François. “LE MERCURE GALANT | Presse18.” Accessed October 26, 2024. <https://dictionnaire-journaux.gazettes18e.fr/journal/0919-le-mercure-galant>.
63. Ovid. “Metamorphoses (Kline) 12, the Ovid Collection, Univ. of Virginia E-Text Center.” Accessed October 26, 2024. <https://ovid.lib.virginia.edu/trans/Metamorph12.htm>.
64. Perrault, Charles. *Histoire de la Marquise-Marquis de Banneville par Charles Perrault et l’abbé de Choisy*. Paris: La Centaine, 1928.
65. Perrault, Charles (1628-1703) Auteur du texte. “Contes ou Histoires du temps passé - Les Contes de ma Mère l’Oye,” 1879. https://fr.wikisource.org/wiki/Les_Contes_de_ma_m%C3%A8re_l%E2%80%99Oye_ava nt_Perrault.
66. ———. “La Belle au bois dormant.” *Mercuré galant*, February 1, 1696. Gallica.
67. Perrault, Charles (1628-1703) Auteur du texte, and Charles-Athanase (1771-1852) Auteur du texte Walckenaer. *Les Contes de Perrault / Précédés d’une Notice Sur l’auteur Par Le Bibliophile Jacob. et Suivis d’une Dissertation Sur Les Contes de Fées / Par Le Baron Walckenaer,...*, 1861. <https://gallica.bnf.fr/ark:/12148/bpt6k6566925v>.

68. Perrault, Charles, Choisy, and Marie-Jeanne L'Héritier de Villandon. *Histoire de la marquise-marquis de Banneville*. Edited by Joan E. DeJean. Texts and translations. New York, NY: The Modern Language Association of America, 2004.
69. ———. *The story of the Marquise-Marquis de Banneville*. Edited by Joan E. DeJean. Texts and translations. New York, NY: The Modern Language Association of America, 2004. <https://archive.org/details/storyofmarquisem00char/>.
70. Raynard, Sophie. *La seconde préciosité: floraison des conteuses de 1690 à 1756*. Biblio 17. Tübingen: G. Narr, 2002. <http://catalogue.bnf.fr/ark:/12148/cb388401677>.
71. READ-COOP. "Transkribus Print Multi-Language." Accessed October 28, 2024. <https://readcoop.eu/model/transkribus-print-multi-language-dutch-german-english-finnish-french-swedish-etc/>.
72. Roche-Mazon, Jeanne. "An unexpected collaboration in the 17th century: the abbé de Choisy and Charles Perrault." *Mercure de France*, February 1, 1928. Gallica.
73. "Sadgrl.Online." Accessed October 30, 2024. <https://goblin-heart.net/sadgrl/webmastery/downloads/tiledbgs>.
74. Scott, P. "Authenticity and Textual Transvestism in the Memoirs of the Abbe de Choisy." *French Studies* 69, no. 1 (January 1, 2015): 14–29. <https://doi.org/10.1093/fs/knu230>.
75. Seifert, Lewis C. "Introduction: Queer(Ing) Fairy Tales." *Marvels & Tales* 29, no. 1 (2015): 15–20. <https://doi.org/10.13110/marvelstales.29.1.0015>.
76. Steinberger, Déborah. *Women's Stories in "Le Mercure Galant" (1672-1710): Feminine Fictions in an Early French Periodical*. Gendering the Late Medieval and Early Modern World 26. Amsterdam: Amsterdam University Press, 2024.
77. Storer, Mary Elizabeth. *Un épisode littéraire de la fin du 17e siècle.: La mode des contes de fées (1685-1700)*. Paris: Champion, 1928.
78. "SurLaLune Fairy Tales: Cenerentola: Giambattista Basile's Il Pentamerone (Story of Stories) 1911 Version," November 23, 2019. <https://web.archive.org/web/20191123051321/http://www.surlalunefairytales.com/pentamerone/6cenerentola1911.html>.
79. Taylor, Helena. "The Quarrel of the Ancients and Moderns." *French Studies: A Quarterly Review* 74, no. 4 (2020): 605–20.

80. Taylor, Joanna. "The Evolution of the Romantic Fairy." *Romantic Textualities* (blog), June 11, 2014. <https://www.romtext.org.uk/the-evolution-of-the-romantic-fairy/>.
81. "Team | Ghent Centre for Digital Humanities." Accessed October 28, 2024. <https://www.ghentcdh.ugent.be/team>.
82. USA, Bryan Allison from Las Vegas, NV. *Photo: Un Joli Chaton Populaire Avec Les Touristes Au Colisée de Rome*. May 20, 2005. Colosseum cat. <https://commons.wikimedia.org/wiki/File:Kitty-popular-with-the-tourists-in-colosseum-Rome-byBryanAllison.jpg>.
83. Warner, Marina. *Once upon a Time: A Short History of Fairy Tale*. First edition. Oxford, United Kingdom: Oxford University Press, 2014. <http://catdir.loc.gov/catdir/enhancements/fy1510/2015413705-t.html>.
84. Weatherley, Gillian Avril. "Avatars of Gendered Societal Constructs in Seventeenth-Century Contes de Fées," 2014. <https://core.ac.uk/download/pdf/213406271.pdf>.
85. Wood, Thomas. "Fast Stylometry," May 1, 2024. <https://doi.org/10.5281/zenodo.11096941>.
86. "WorldCat.Org." Accessed October 27, 2024. <https://search.worldcat.org>.
87. "Zeob: Create Fake Instagram Posts With Zeob's Post Generator Tool," July 16, 2022. <https://zeob.com/generate-instagram-post/>.
88. Zipes, Jack. *The Oxford Companion to Fairy Tales*. Accessed October 27, 2024. https://pubhtml5.com/pqqj/yjyv/The_Oxford_Companion_to_Fairy_Tales/.
89. Zipes, Jack David. *The Great Fairy Tale Tradition : From Straparola and Basile to the Brothers Grimm : Texts, Criticism*. New York : W.W. Norton, 2001. <http://archive.org/details/greatfairytaletr00zipe>.