

# Colette, Curnonsky, and the Willy workshop. Assessing relative contributions and influences beyond “collaborative authorship”

**Cafiero, Florian**

florian.cafiero@sciencespo.fr  
Sciences Po, France

**Puren, Marie**

marie.puren@epita.fr  
EPITA, France; Centre Jean-Mabillon, France

What was the influence on the French writer Colette of her first husband Willy and his literary workshop? Her opinion varied greatly, depending on the time she spoke. When asked just after the publication of the *Claudine*, a series of books published from 1900 to 1903 initially under the sole name of her husband Willy, Colette admitted her contribution with a hint of a smile: “[Willy] crie si obstinément ma participation [aux *Claudine*] que je dois m’incliner” (“[Willy] so obstinately cries out for my participation [in the *Claudine*] that I must bow down”). But in 1936 in her autobiographical work *Mes Apprentissages*, and until her death in 1954, Colette changed her tone and insisted on Willy’s minor role in the writing of these books, until her first husband’s name disappeared from the cover of the novels. When Colette was asked if Willy helped her to write these texts, she replied in 1949: “Plutôt par des indications, mais ça ne peut s’appeler une aide...” (“Rather by indications, but it cannot be called a help ...”). Which Colette should we believe?

In this paper, we analyze texts that we have occluded and corrected. Using artificial intelligence profiling techniques, we show that *Claudine*’s texts at school bear the influence of Willy and his workshop.

## Claudine at school: a collaborative work?

Sidonie Gabrielle Colette, known as Colette (1873-1954), took her first steps on the French literary scene in the shadow of her first husband Henry-Gauthier Villars (1859-1931), better known under his pseudonym Willy. From 1900 to 1903, four novels belonging to the *Claudine* series were published and signed by Willy: *Claudine à l’école* (1900), *Claudine à Paris* (1901), *Claudine en ménage* (1903), *Claudine s’en va* (1903). Willy never hid the fact that Colette had participated in the writing of these books, as Colette acknowledged in *Mes Apprentissages*. The use of his name is explained both by commercial reasons (Willy was a successful author) and by social reasons: as a woman in 1900, Colette might have encountered many difficulties in publishing a novel under her own name. Willy’s methods of literary production are well known to his Parisian colleagues and friends. In 1900, Willy was known to employ numerous “secretaries” or “collaborators” to write “his”

books. Colette describes their apartment as a “literary workshop”, where several ghostwriters wrote novels, meticulously revised, rewritten and commented by Willy. During his literary career, it is estimated that about fifty collaborators were employed by Willy. The draft manuscript of *Claudine à l’école* would also have been “processed” by this workshop; and after being revised by Willy, it would have been reduced by half (Caradec 2004).

Nevertheless, we do not know exactly which part of the text was written by Colette and which part was rewritten (or even written) by Willy or his collaborators. To what extent did Willy influence the first literary texts written by Colette? What role did he really play in the writing of these novels?

## A methodological challenge

Analyzing the relative contributions of Colette, Willy and his secretaries poses many methodological problems. First, texts that can be attributed with certainty to Willy alone are very rare - even more so if we restrict ourselves to the novelistic genre. Finding texts written in isolation by each of the participants in Willy’s workshop is also a complex task. These writers are numerous, not all of them have pursued a literary career, and some of them have probably continued to write in collaboration with other authors. This makes it very difficult to treat the *Claudines* as other cases of collaborative authority (Plechac 2019; Cafiero / Camps 2021): training an artificial intelligence to recognize each candidate’s style and then applying a rolling stylometry algorithm (Eder 2016) is a short-term challenge, given the lack of reliable texts available.

Moreover, if Colette is telling the truth, Willy’s main influence on these texts would come from comments and suggestions made on a text written and then corrected by herself. If this is indeed the case, a classical stylometric approach would fail: it would only detect Colette’s hand, but not her influences. (Cafiero / Camps 2019).

## Profiling

If we cannot easily try to recognize Willy’s writing in Colette’s text, we can try to identify his influence in other ways. The reasoning is simple: compare texts written by authors in Willy’s workshop to those they wrote alone, outside the workshop, at the closest possible date to their release. The difference between these texts? To a first approximation, it is what we will call here the “Willy effect”. It includes Willy’s corrections, advice or writing passages, but is not reduced to them. Others than him have been able to help their colleagues to make their manuscript closer to Willy’s expectations - by proofreading, giving advice, etc.

To measure this effect, we employ profiling logic (Bevendorf et al. 2021; Daelemans et al. 2019; Mikros 2017). We train an artificial intelligence (Support Vector Machine) to distinguish between texts written by certain authors in collaboration with Willy and texts written by the same authors outside of Willy’s workshop. The training is done on a corpus of novels by Paul Acker, Maurice Edmond Sailland (known as Curnonsky), Jeanne Marais, Jean de Tinan, Léo Trézenik and Pierre Veber. The SVM takes as input a combination of the most frequent character bi-grams, character tri-grams, words, and word bi-grams. The optimal parameters are deduced by cross-validation. A polynomial kernel is adopted, which gives better performance here than different linear kernels. Low variance features are removed to reduce the compu-

tation time. The performance of this profiling method is satisfactory (F-score: 0.89).

The texts to be evaluated (Colette's novels) are then cut into sections of a thousand words, a value for which performances were deemed a satisfactory compromise between performance and granularity (best F-Score: 0.93 for 3000 words and above). For each section, the SVM is asked if the text seems to be influenced by a Willy effect or not.

## Results

The influence of Willy's workshop seems to be obvious on the *Claudines*. Colette's first novel, *Claudine à l'école*, seems particularly sensitive to it. But all the novels of the series seem to bear the trace of it.

The proportion of texts influenced by Willy and his workshop seems to diminish over time, without disappearing. Is this just an artifact? Is it, on the contrary, the trace of an impact of this workshop on Colette's later writing, which would have remained marked by this period? Did members of the workshop, such as her friend Curnonsky, advise or reread Colette's passages after she left Willy's workshop? It is impossible to decide in the present state of the study. But these are all avenues that we will have to explore next and present at the conference.

	<i>Claudine à l'école</i>	<i>Claudine à Paris</i>	<i>Claudine en ménage</i>	<i>Claudine s'en va</i>	<i>Le blé en herbe</i> (1923)	<i>La Chatte</i> (1933)	Julie de Carnéilhan (1941)
Proportion of segments detected as "in collaboration with Willy and his workshop"	0,45	0,45	0,38	0,37	0,13	0,27	0,24

## Bibliography

**Bevendorff, Janek / Chulvi, Berta / Peña Sarracén, Gretel Liz De La / Kestemont Mike / Manjavacas Enrique / Markov Ilija / Mayerl Maximilian / Potthast Martin / Rangel Francisco / Rosso Paolo / Stamatatos Efstathios / Stein Benno / Wiegmann Matti / Wolska Magdalena / Zangerle Eva** (2021): "Overview of PAN 2021: Authorship Verification, Profiling Hate Speech Spreaders on Twitter, and Style Change Detection", in: *International Conference of the Cross-Language Evaluation Forum for European Languages*. Cham: Springer 419-431 [https://dbis.uibk.ac.at/sites/default/files/2021-09/clef21\\_pan\\_overview.pdf](https://dbis.uibk.ac.at/sites/default/files/2021-09/clef21_pan_overview.pdf) [01.05.2023].

**Caffero, Florian / Camps, Jean-Baptiste** (2019): "Why Molière most likely did write his plays", in: *Science advances* 5, 11: eaax5489 <https://www.science.org/doi/10.1126/sciadv.aax5489> [01.05.2023].

**Caffero, Florian / Camps, Jean-Baptiste** (2021): "'Psyché' as a Rosetta Stone? Assessing Collaborative Authorship in the French 17th Century Theatre", in: Ehrmann, Maud / Karsdorp, Folger / Wevers, Melvin / Andrews, Tara Lee / Burghardt, Manuel / Kestemont, Mike / Manjavacas, Enrique / Piotrowski, Michael / Zundert, Joris van (eds.): *Proceedings of the Conference on Computational Humanities Research 2021*. Amsterdam: CEUR Workshop Proceedings 1613: 377-391 [https://ceur-ws.org/Vol-2989/long\\_paper51.pdf](https://ceur-ws.org/Vol-2989/long_paper51.pdf) [01.05.2023].

**Caradec, François** (2004): *Willy: le père des Claudine*. Paris: Fayard.

**Daelemans, Walter / Kestemont, Mike, Manjavacas, Enrique / Potthast, Martin / Rangel, Francisco / Rosso, Paolo / Specht, Günther / Stamatatos, Efstathios / Stein, Benno / Tschuggnall, Michael / Wiegmann, Matti / Zangerle, Eva** (2019): "Overview of PAN 2019: bots and gender profiling, celebrity profiling, cross-domain authorship attribution and style change detection", In: *International conference of the cross-language evaluation forum for european languages*. Springer, Cham 11696: 402-416 [https://doi.org/10.1007/978-3-030-28577-7\\_30](https://doi.org/10.1007/978-3-030-28577-7_30) [01.05.2023].

**Eder, Maciej** (2016): "Rolling stylometry", in: *Digital Scholarship in the Humanities* 31, 3: 457-469 <https://doi.org/10.1093/llc/fqv010> [01.05.2023].

**Mikros, George K.** (2018): "Blended Authorship Attribution: Unmasking Elena Ferrante Combining Different Author Profiling Methods", in: *Drawing Elena Ferrante's Profile*. Padova: Padova University Press 85-96 <http://www.padovauniversitypress.it/system/files/preview/9788869381300.pdf> [01.05.2023].

**Plecháč, Petr** (2021): "Relative contributions of Shakespeare and Fletcher in Henry VIII: An analysis based on most frequent words and most frequent rhythmic patterns", in: *Digital Scholarship in the Humanities* 36, 2: 430-438 <https://doi.org/10.1093/llc/fqaa032> [01.05.2023].