CORA Cork Open Research Archive
Cartlann Taighde Oscailte Chorcaí

Title	Who wrote Els Quaderns d'en Marc? A stylometric inquiry into twentieth-century Catalan popular literature
Authors	Nilsson-Fernàndez, Pedro
Publication date	2025-12-08
Original Citation	Nilsson-Fernàndez, P 2025, 'Who wrote Els Quaderns d'en Marc? A stylometric inquiry into twentieth-century Catalan popular literature', Digital Scholarship in the Humanities, vol. 41, no. 1, pp. 320-332. https://doi.org/10.1093/llc/fqaf096
Type of publication	Article (peer-reviewed)
Link to publisher's version	https://doi.org/10.1093/llc/fqaf096 - 10.1093/llc/fqaf096
Rights	© 2025, the Author(s). Published by Oxford University Press on behalf of EADH. This is an Open Access article distributed under the terms of the Creative Commons Attribution-NonCommercial License (https://creativecommons.org/licenses/by-nc/4.0/), which permits non-commercial re-use, distribution, and reproduction in any medium, provided the original work is properly cited. - https://creativecommons.org/licenses/by-nc/4.0/
Download date	2026-11-16 07:01:42
Item downloaded from	https://hdl.handle.net/10468/19494

UCC
University College Cork, Ireland
Coláiste na hOllscoile Corcaigh

Digital Scholarship in the Humanities, 2026, **41**, 320–332
https://doi.org/10.1093/llc/fqaf096
Advance access publication 8 December 2025
Short Paper

OXFORD

Who wrote *Els quaderns d'en Marc*? A stylometric inquiry into twentieth-century Catalan popular literature

Pedro Nilsson-Fernàndez[1,*]

[1]School of Languages, Literatures and Cultures, University College Cork, Cork T12 ND89, Ireland
*Corresponding author. School of Languages, Literatures and Cultures, University College Cork, Cork T12 ND89, Ireland. E-mail: pfernandez@ucc.ie

Abstract

The anonymous erotic Catalan trilogy 'Els quaderns d'en Marc' (1984–85) remains one of the most intriguing mysteries in twentieth-century Catalan popular imagination. While many in the literary circles attribute the work to the Catalan polygraph writer Manuel de Pedrolo (1918–90), there has never been a clear confirmation from the author himself, or any significant study dedicated to exploring the subject in detail. With the use of stylometry, in particular the Stylo package for R, the main objective of this short paper is to look at the authorial fingerprint in the first three volumes of 'Els quaderns d'en Marc' (1984–85), in the first instance, and at the sequel 'Els Quaderns de la dona d'en Marc' (1985), in the second. The aim is to examine whether this methodology can validate or demystify some of the existing theories surrounding these anonymous works' authorship. The main contribution of this analysis has been to confirm—by the use of three separate approaches to stylometry as recommended by Juola—the high degree of similarity between the authorial fingerprint of Manuel de Pedrolo and that of the anonymous author of volumes I to III of this saga. As part of this authorship attribution exercise, a second finding has emerged; in the case of volume IV of this saga, 'Els quaderns de la dona d'en Marc' (1985), the writer Manuel Joan i Arinyó, who was initially a suspected candidate, has been identified as a strong potential author of this volume.

Keywords: quaderns; stylometric; digital; stylometry; delta; Catalan; authorship; stylo; Pedrolo.

1. Introduction

On 23 November 1983, Enric Borràs, director of the small Barcelona-based Catalan publishing house El Llamp, received a package in the post containing a mysterious typescript. Upon examining the text, Borràs learned that the work was an anonymous erotic trilogy narrating the sexual encounters between the male protagonist Marc and a number of Catalan women. The text was accompanied by a short but clear set of instructions for publication. The text would later be included as an introductory chapter to the published version of the work, a three-volume novel titled *Els quaderns d'en Marc* (1984a, 1984b, 1985a). The work still stands as the most notorious erotic work written in Catalan, and its authorship has been unofficially attributed to Manuel de Pedrolo i Molina (1918–90). The key role this writer played in the production and promotion of popular genres in Catalan language,[1] together with some circumstantial evidence surrounding the work's original typescript, suggests that Pedrolo could very likely be the author. Nonetheless, the writer publicly denied any involvement in the authorship of this work (Miquel 2015), and even though there is a variety of opinions voiced by members of the Catalan literary world,[2] these mentions, as we will see in later sections, are mere speculation; as Bel Zaballa notes in the author's biography, 'all these suspicions are mere clues and there is no conclusive proof' (2018). Moreover, discussion of the authorship of this work has never been the sole focus of any extant piece of research; the question of who wrote the text is mentioned only in passing, while the attention falls upon other aspects of the work, such as its originality, its literary quality, or the impact it had in enabling a wave of publications in the genre during the 1980s. Thus, this inquiry is not only a response to that gap, but also the first to use computational tools as a methodology to address authorship in the context of Catalan literature.

With the use of stylometry, in particular the Stylo package for R (Eder et al. 2016), the main aim of this

© The Author(s) 2025. Published by Oxford University Press on behalf of EADH.
This is an Open Access article distributed under the terms of the Creative Commons Attribution-NonCommercial License (https://creativecommons.org/licenses/by-nc/4.0/), which permits non-commercial re-use, distribution, and reproduction in any medium, provided the original work is properly cited. For commercial re-use, please contact reprints@oup.com for reprints and translation rights for reprints. All other permissions can be obtained through our RightsLink service via the Permissions link on the article page on our site—for further information please contact journals.permissions@oup.com.

Downloaded from academic.oup.com/dsh/article/41/1/320/8374195 by University College Cork user on 22 October 2026

paper is to look at the authorial fingerprint in the first three volumes of *Els quaderns d'en Marc*[3] (1984a, 1984b, 1985a), in the first instance, and at the sequel *Els Quaderns de la dona d'en Marc* (1985b), in the second. The aim is to examine whether this methodology can validate or demystify some of the existing theories surrounding this anonymous work's authorship. As this article will set out in detail, we differentiate this first trilogy from the later sequels to this work: *Els Quaderns de la dona d'en Marc* (1985b), *Mateu, el fill d'en Marc* (1987) and *Diari de campanya (El fill d'en Marc a Prada)* (1987). We know that the last two works are signed by Valencian writer Manuel Joan i Arinyó (1956). When it comes to *Els Quaderns de la dona d'en Marc* (1985b), also anonymous, its absence from the first trilogy's introductory chapter indicates a certain fragmentation from the original three volumes, which justifies a separate analysis.

The importance of shedding some light on *Els quaderns*' authorship goes beyond solving a mere matter of attribution; were Pedrolo to be the author of this work—'the most controversial and commented erotic Catalan novel' (Vilaweb)—it would constitute a very meaningful addition to the author's already crucial contribution to the twentieth-century Catalan popular literature canon. While Pedrolo is known to be one of the most influential Catalan writers of sci-fi and noir in the twentieth century, credit for writing the most famous Catalan erotic work of his time, would further consolidate his wider literary reputation in the context of Catalan popular literature.

2. A short note on stylometry

Stylometry, the methodology we employ in this article, can be broadly defined as the statistical analysis of literary style, which relies on examining how frequently words appear in a given text to determine aspects such as its authorial fingerprint. Drawing on early twentieth-century studies such as George Kingsley Zipf's Principle of Relative Frequency, which claims that '[t]he accent or degree of conspicuousness of any word, […] is inversely proportionate to the relative frequency of that word', (Zipf 1932) stylometric methodologies evolved through the later twentieth century with key contributions such as Frederick Mosteller and David Wallace's study of eighteenth-century *Federalist Papers* in 1964,[4] Bradley Efron and Ronald Thisted's approach to a Shakespearean authorship of poetry problem in 1987,[5] and were regenerated by John Burrows in the 1980s,[6] who introduced the use of larger datasets in his approach. A key contribution of the so-called 'Burrows's method' is the general consensus that the most frequent words in a given text—function words—are those to be considered in an authorship inquiry, rather than the rarer—non-function words—that scholars such as Efron and Thisted (1976) focused on in their earlier studies. Nowadays, the aid of powerful computational approaches to these theories, in the form of easily accessible open-source tools such as Stylo package for R,[7] which is used in this study, we can apply these theories with relative ease. Thus, when we refer to the concept of 'authorial fingerprint' in the following sections, we are referring to an authorial profile extracted from identifying patterns of usage of the most frequent words by an author in his signed texts, in this case, Manuel de Pedrolo. In the same way, we aim at identifying the profile of the anonymous author of *Els quaderns*, in an attempt at comparing it to that of Manuel de Pedrolo or any other potential candidates.

3. The original typescript and its introductory chapter

There are two pieces of information in the above-mentioned introductory chapter that are key to this analysis. First, the chapter introduces the trope of a 'rewritten manuscript', which alludes slightly to Manuel de Pedrolo's practices, and second, it informs us of the three-volume structure of the cycle. The anonymous writer begins this chapter by identifying himself as an established Catalan author[8] who, sometime in Autumn of 1968, received the visit of a seemingly inoffensive reader who wanted to get a few copies of his books signed. During a second visit, this admirer confessed that the real reason for contacting him was to ask for his help in publishing the original manuscript of *Els quaderns*, a handwritten text containing his personal account of erotic experiences. The anonymous writer agreed, with the only condition of being able to rewrite a work that he considers, though interesting in its content, to be written by someone who 'is rather clumsy with the pen in his hand'[9] (Anon 1984a). This rewriting exercise can be observed in other works by Manuel de Pedrolo,[10] but it is not sufficient evidence to confirm his authorship.

The second piece of information in this introduction is its very well-defined structure: the anonymous author states that *Els quaderns* was divided into three volumes. This detail is quite important, for it differentiates *Els quaderns de la dona d'en Marc* from the previous three novels: *Els quaderns d'en Marc I*, *II*, and *III*. While this distinction does not necessarily confirm or disprove the idea of all four works sharing author, it is important to note that the introductory note clearly differentiates these first three volumes from the rest of the sequels, which receive no mention at all, and were clearly not part of the writer's plan at the time.

Beside these two first points, there are several others that build some circumstantial evidence and that are

Downloaded from academic.oup.com/dsh/article/41/1/320/8374195 by University College Cork user on 22 October 2026

better discussed by reading them in their original sources; in the next section, the reader will see what a number of Catalan scholars, writers, and intellectuals have said about this literary mystery throughout the years.

4. The voices

There is very little literature that mentions Manuel de Pedrolo's possible involvement in *Els quaderns*, particularly considering how widespread the assumption of his authorship is among Catalan scholarship and readership. Pedrolo's established position as a Catalan popular fiction writer in the 1980s and the fact that this popularity relied heavily on two works with a pronounced presence of sexuality/eroticism—his crime fiction novel *Joc brut* (1965) and the sci-fi bestseller *Mecanoscrit del segon origen* (1974)—may have contributed to spreading the rumour of him being the author of *Els quaderns*. What is more, a signed erotic novel by the author, *Obres púbiques* (1991), was published posthumously but written in 1961, clearly indicating that the writer had an interest in cultivating this genre from an early stage in his career.

The first and quite relevant testimonial is that of Enric Borràs (2005), the editor of the trilogy already mentioned in the introduction. At the time of receiving the anonymous text, Borràs had only recently published a collection of Manuel de Pedrolo's political articles, *Cròniques colonials* (1982). This detail is important, as the editor suggests a connection between the paper used to type the anonymous work and previous materials submitted by Manuel de Pedrolo. As Borràs writes in 2005, Pedrolo had mentioned to him that, since losing the original typescript of a novel in 1977, he had started using a particular brand of tracing paper that produced two additional copies of each typewritten page (Miquel 2015). This exact practice is also described in detail in the introductory chapter, where the anonymous author talks about the rewriting process of Marc's manuscript. He mentions a 'rooted practice of keeping extra copies since, years ago, I had lost the only copy of a work I have never rewritten again' (Anon 1984a). While we cannot rule out the possibility that any other Catalan writer at the time may have used that same type of tracing paper, Borràs had grounds to suspect that Pedrolo was the author, to the extent that he confronted him on one occasion and asked if he had indeed written it. Pedrolo clearly denied any involvement, and even asked Borràs if he could read the attached letter and help guess who the author might be (Miquel 2015).

Another important contribution to this debate is that of Joaquim Carbó, who based the plot of his novel *S'ha acabat el bròquil* (1987) on an investigation into *Els quaderns*' authorship. Even through fiction, Carbó puts Manuel de Pedrolo on the spot; the two main characters of this novel are two lovers—he an aspiring writer, and she a scholar who specializes in Pedrolo—that manage to enter Pedrolo's study while he is not at home, in a failed attempt to discover the original manuscripts of *Els quaderns*. However, the relevance of this novel to the question at hand goes beyond its pages; sometime in mid-1987, Joaquim Carbó visited Pedrolo to ask him about his opinion of the novel *S'ha acabat el bròquil*. We find a description of this encounter in a letter Carbó writes to another Catalan writer, Jordi Arbonès, on November 8th of the same year (Carbó 2005). On that occasion, Pedrolo had a very difficult time telling Carbó that he had not liked the treatment of sexuality in the novel. Besides Pedrolo's clear avoidance of the topic of authorship, Joaquim Carbó's comments suggest that he himself did not believe Pedrolo to have written *Els quaderns:*

> All of this makes me think that, as Albanell was telling me the other day, it looks as if the supposition that Pedrolo is the author of *Els quaderns d'en Marc* is as false as dice, when everything had seemed to point in that direction. (Bacardí 428)

To whom Arbonès, in turn, replies on 18 February 1988:

> I also had my reservations about the Pedrolian paternity of *Els quaderns d'en Marc*. In my book I already pointed out that, surprisingly, Pedrolo never used vulgar words to refer to sexual actions or organs; if anything, he used the French words instead. In conclusion, he is more of a moralist than he seems. (Bacardí 433–34)

Thus, the most important detail we learn from this discussion around *S'ha acabat el bròquil* is that Pep Albanell, Joaquim Carbó, and Jordi Arbonès—three Catalan writers in Pedrolo's close circle—were under the impression that *Els quaderns d'en Marc* were not written by him. This is, at least, during the late 1980s. Twenty-five years later, however, one of these three, Joaquim Carbó, will look at the parallelism between Henry Miller's erotic tales *Opus Pistorum*—penned in 1941 and later published as *Under the Roofs of Paris* (1983)—and *Els Quaderns*. He does so by engaging with Catalan writer and translator Jordi Arbonès' comments on the work of both Miller and Pedrolo, two writers he greatly admired.[11] Carbó looks at the fact that both writers firmly denied the authorship of these erotic works, and notes how both texts include an epilogue related to authorship. In Miller's case, however, there is no room for doubt when it comes to authorship, as the question was answered years after

Downloaded from academic.oup.com/dsh/article/41/1/320/8374195 by University College Cork user on 22 October 2026

Miller passed away. In the epilogue of *Opus Pistorum*, bookshop owner Milton Luboviski identifies the work as commissioned by him in the 1940s, in an attempt to improve both their meagre economic conditions at the time.

The connection that Carbó does not make between the two works is one that can be drawn through 'the acknowledgement of the figure of Manuel de Pedrolo as a relentless, committed, combative writer who […] built a lifetime aesthetic project from the conviction that Catalan literature could operate—even from a position of subalternity—at the same level as any other world literature' (Nilsson-Fernàndez 2019). In order to strive for such 'normality,' a big part of Pedrolo's contribution to twentieth-century Catalan literature consisted of bringing foreign models into his own corpus, both through translation and imitation of works of foreign authors such as Miller. Pedrolo translated two works from this American Author: *A Devil in Paradise* (1956) and *My Life and Times* (1975), 'works that were certainly not the most well-known from the author' (Pijuan Vallverdú, 2016). Thus, it would not be a surprise if *Els quaderns* was Pedrolo's Catalan counterpart to Henry Miller's *Opus Pistorum*; in a small literature such as that written in Catalan, nothing says 'normality' better than having a rare anonymous erotic work in its ranks.

In fact, Jordi Arbonès wrote an introductory study of Miller's *Sexus*—a work that he translated into Catalan—pointing to the issue of subalternity permeating all aspects of Catalan literature and culture. Arbonès finds the minoritised status of his literature essential to understand why no other studies have been dedicated to the authorship of *Els quaderns*. Considering the close literary relationship Pedrolo and Arbonès maintained through their letters, his mention of *Els quaderns* in a study of an American erotic novel is particularly significant:

> If these interesting, entertaining episodes had appeared in a different literary atmosphere, even a slight intervention by a curious reporter or literary critic into the connection with Pedrolo would have caused some major drama. This, however, is something unthinkable in this sort of cemetery in which our writers are only read by each other and self-proclaim themselves unstoppable while they know that no first edition of their work will ever sell more than fifteen hundred copies. (Miller, 1992)

It could just be a mere coincidence, but the date in which Miller's heirs won the legal case that allowed them to monetize the previously anonymous *Opus Pistorum* is precisely 1983, the same year that *Els quaderns* appear at El llamp publishing house. It is then a possibility that the author, regardless of his identity, was aware of Henry Miller's work and his relationship with the bookshop owner. What is certainly true is that the anonymous writer is an informed connoisseur of erotica; the introduction of *Els quaderns* mentions works such as the also anonymous *My Secret Life* (1966), Frank Harris' *My Life and Loves*, and Simon Forman's papers (Anon 1984a).

A long time after Pedrolo's passing in 1990, reactions to the work and its authorship have continued to appear, particularly online. On the occasion of the 2001 re-edition of the work by publishing house La Magrana, Joan Josep Isern wrote a review of the work claiming that its literary quality had allowed the work to read well after fifteen years since its original publication. Again, in passing, he refers to the collective consensus reached by twentieth-century Catalan readers, who attribute the work to Pedrolo, as 'incredibly surprising in a country where everything is done *en petit comité*' (2001), as well as irrational, as the author never admitted to being the writer.

In later years, literary blogs specialised in Catalan literature have intermittently kept the debate going, while giving visibility to the work. This is the case with writer Iolanda Pàmies, who noted that 'the beauty of some fragments, the descriptions that introduce the characters and their stories to us, cannot have been penned by any clumsy amateur' (2010). She goes on to mention Pedrolo, Joan Fuster, Pere Quart or Vicent Estellés as possible authors, and traces the opening sentence 'Catalonia will be erotic or it will not be at all' to Catalan writer Joan Fuster's appropriation of writer and bishop Torras i Bages' words 'Catalonia will be Christian or it will not be at all'. She argues that Fuster uses this sentence to fit his own political agenda with his own take reading 'Valencia county will be leftist or it will not be at all'. The fact that this was originally a religious quote also fits with Pedrolo's reinterpretation of religious themes in many of his works.[12]

In another blog entry from 2016, Francesc Llaverias gives us what is perhaps the most detailed account of identifying Manuel de Pedrolo's style in the novel through close reading:

> […] I read the book and the only thing I can say is that Pedrolo is the author, beyond any reasonable doubt. The style is his, the language is his, and the soul of the characters is Pedrolian. Almost too much. If Pedrolo wrote the book and did want to publish it anonymously, he should not have been so recognisable. It is absurd, for Pedrolo was completely capable of masking his style—he was a more than experienced writer—so why didn't he try and hide it? Maybe because it wasn't him? Is it perhaps a different author imitating Pedrolo's style? (Llaverias 2016)

Downloaded from academic.oup.com/dsh/article/41/1/320/8374195 by University College Cork user on 22 October 2026

All that can be ascertained from analysing traditional discourses around *Els quaderns d'en Marc* is that, while suspicions frequently point towards Pedrolo, there is still no certainty as to its true authorship. In the following sections, we will attempt to add another angle to the discussion by employing computer-assisted techniques to explore the question of authorship.

5. The corpus: sources and treatment

The corpus employed for this analysis consists of the anonymous texts and a number of samples from potential candidates. As explained before, the three first anonymous texts we are considering are volumes I, II and III of *Els quaderns d'en Marc*, which will be treated as a unit when performing the analysis.

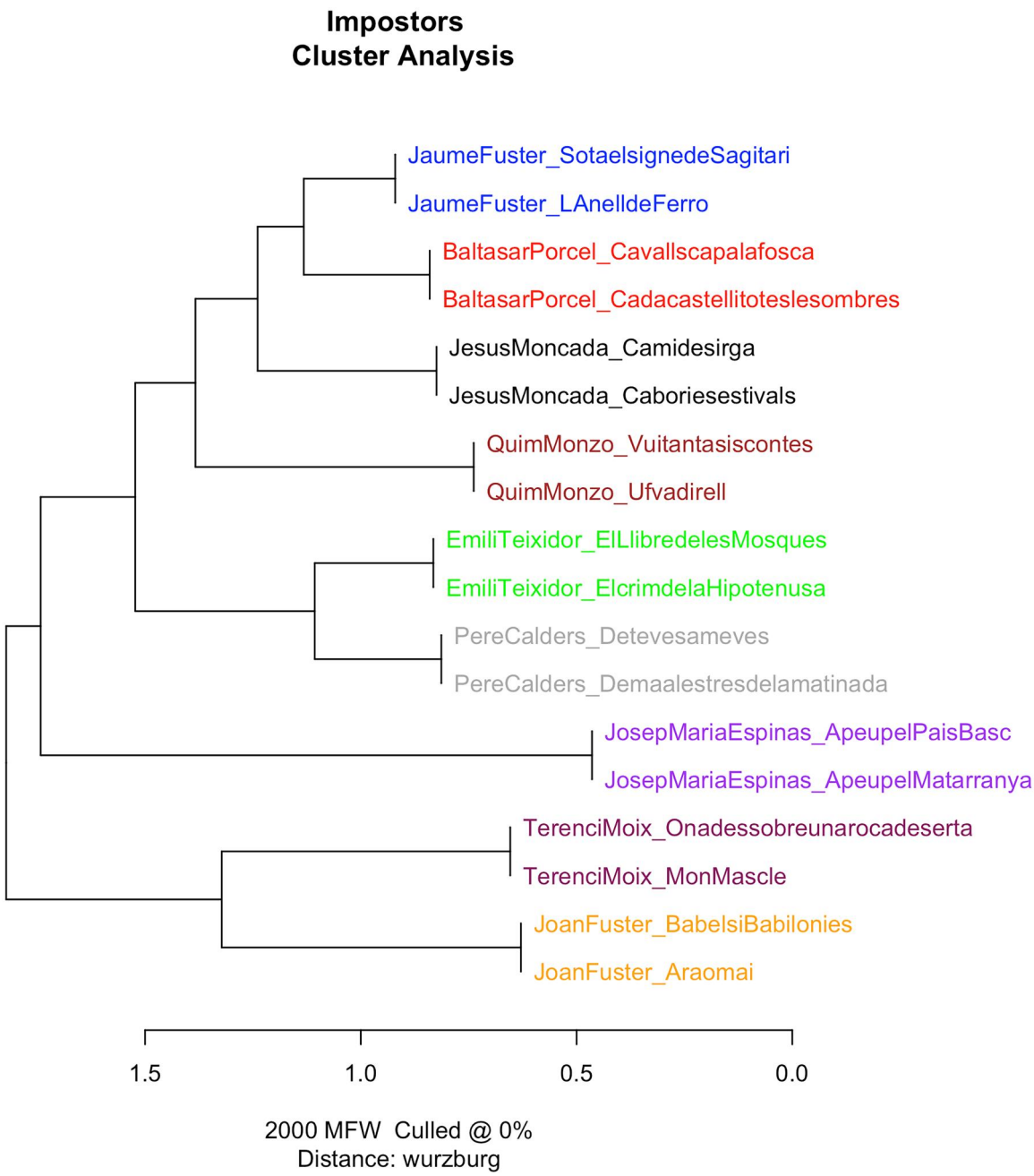


Figure 1. Cluster analysis of novels by nine Catalan authors contemporary to *Els quaderns* (Cosine Delta, 2000 MFW).

Downloaded from academic.oup.com/dsh/article/41/1/320/8374195 by University College Cork user on 22 October 2026

Els quaderns de la dona d'en Marc, the first sequel to the original, is also included as part of this study. When it comes to the potential candidates, we have selected works from Manuel de Pedrolo, as the main suspected writer, and a series of works written by a number of his contemporaries. Some of the writers were mentioned previously in this text as being involved in discussions about authorship. These writers are Joaquim Carbó, Josep Albanell and Manuel Joan i Arinyó. The reason for not including Jordi Arbonès in this discussion is that he was not a novelist; as a translator and critic, the only pieces we could have used would have been his translation of Henry Miller's *Sexus* into Catalan, and an essay about Pedrolo, *Pedrolo contra els límits* (1980). According to extant studies on translation and stylometry (Rybicki 2013;

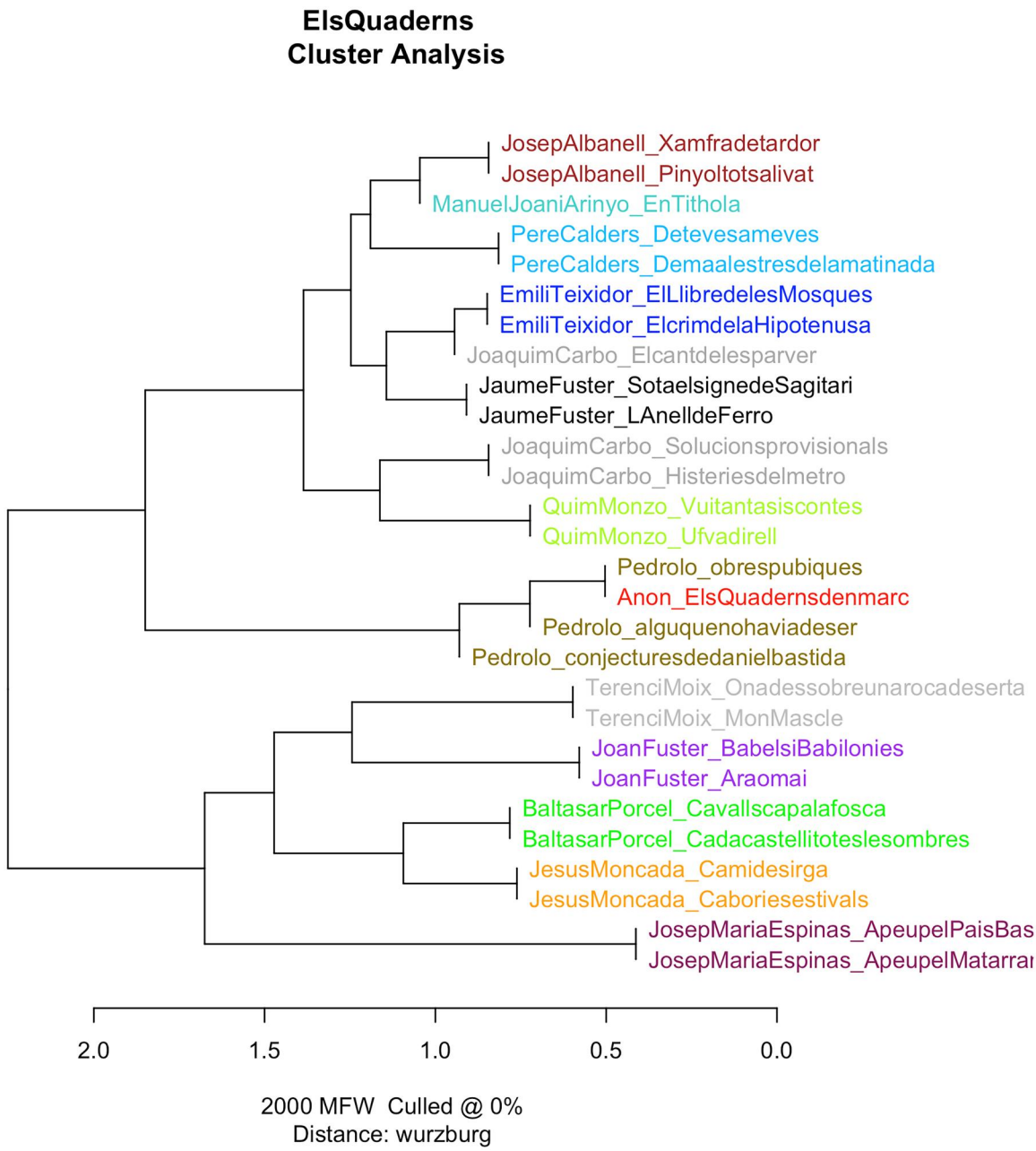


Figure 2. Cluster analysis of the full corpus of potential candidates and *Els quaderns d'en Marc* trilogy (Cosine Delta, 2000 MFW).

Downloaded from academic.oup.com/dsh/article/41/1/320/8374195 by University College Cork user on 22 October 2026

Pagano *et al.* 2024) translated texts tend to keep too many traces from the original writer's authorial imprint to be reliable material for authorship attribution.

Manuel Joan i Arinyó, on the other hand, although absent from discussions around authorship, cannot be disregarded. He is the confirmed author of two of the three sequels of this erotic saga, and also wrote the erotic novel *Stress* (1985), published in the same collection as *Els quaderns*. Later in this study, we will examine this novel in greater detail, further justifying Joan i Arinyó's position as a strong candidate for authorship.

In addition, an auxiliary corpus of eighteen novels by nine male Catalan writers contemporary to Pedrolo and his circle has been curated to serve as *imposters*[13] in our analysis. Two novels from each of the following writers were selected: Baltasar Porcel, Emili Teixidor, Jaume Fuster, Jesús Moncada, Joan Fuster, Josep Maria Espinàs, Pere Calders, Quim Monzó, and Terenci Moix. It is important to note that, unlike in larger literary traditions such as English or Spanish, twentieth-century Catalan literature was sufficiently small to make the identification of potential candidates relatively manageable.[14] The nine writers named above were not chosen randomly from a wider pool. After identifying sufficiently notable Catalan authors of the period and excluding those who wrote exclusively in genres other than prose or in dialects distinct from standard Catalan, the list was reduced to these nine authors.

6. Analyzing the corpus with the Stylo package in R

The first part of the stylometric analysis, conducted with the *Stylo* package for R (Eder, Rybicki, and Kestemont 2016), is informed by the work of Stefan Evert *et al.* (2017). Using the same Stylo package, their study demonstrated that the combination of the Cosine Delta distance and a higher number of Most Frequent Words (MFW)—with results stabilizing at around 2,000 MFW when applied to English, German, and French test corpora—yields optimal outcomes for authorship attribution. The robustness of this approach has been confirmed by later studies (Ochab *et al.* 2019).

Applying the same methodology to our Catalan corpus of eighteen novels (later employed as imposters) we observe in Fig. 1 a clean clustering pattern: each

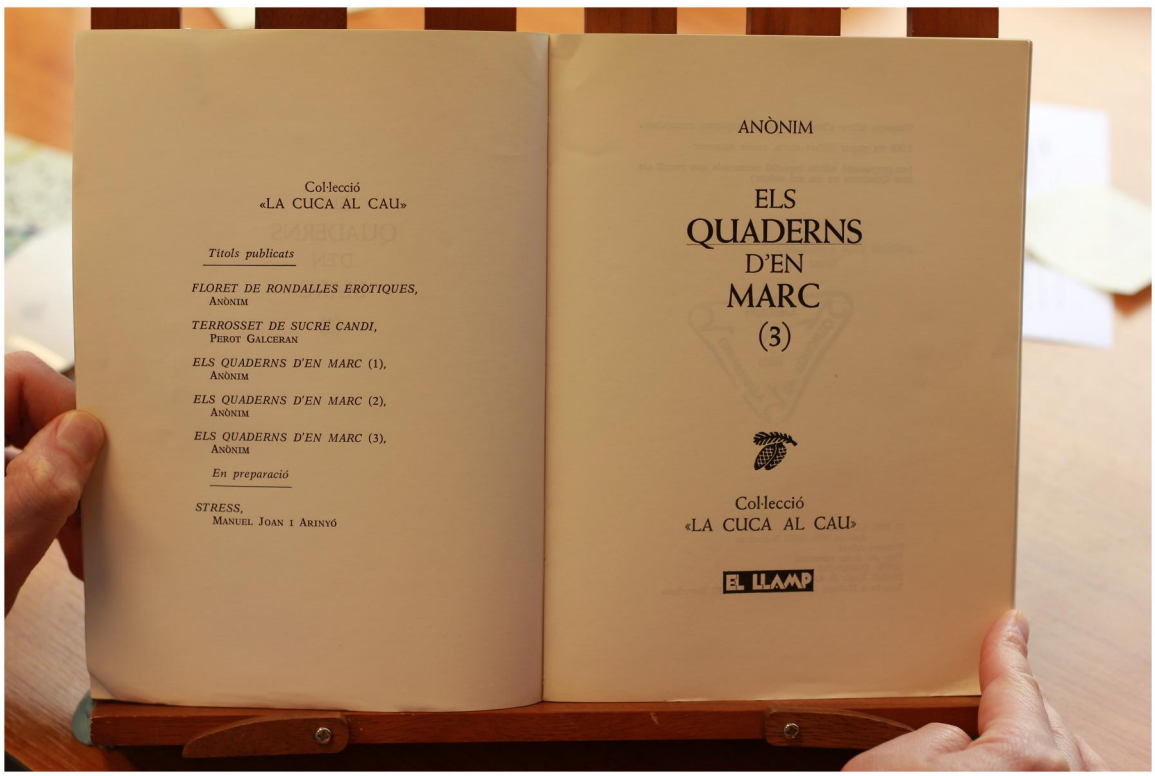


Figure 3. Page two of *Els quaderns d'en Marc* (vol. 3, 1985), noting *Stress* (1985) as 'in preparation' and attributing it to Manuel Joan i Arinyó.

Downloaded from academic.oup.com/dsh/article/41/1/320/8374195 by University College Cork user on 22 October 2026

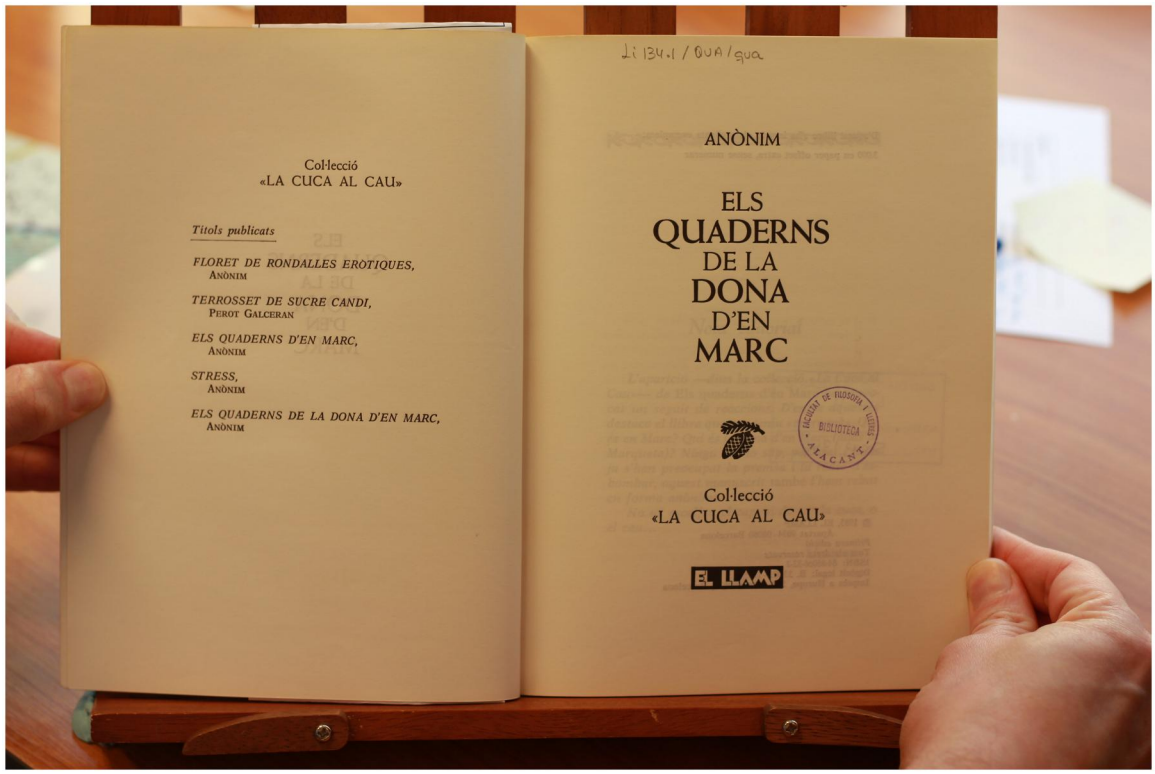


Figure 4. Page two of *Els quaderns de la dona d'en Marc* (1985), referencing *Stress* (1985) as already published in the collection, but attributed here to an anonymous author.

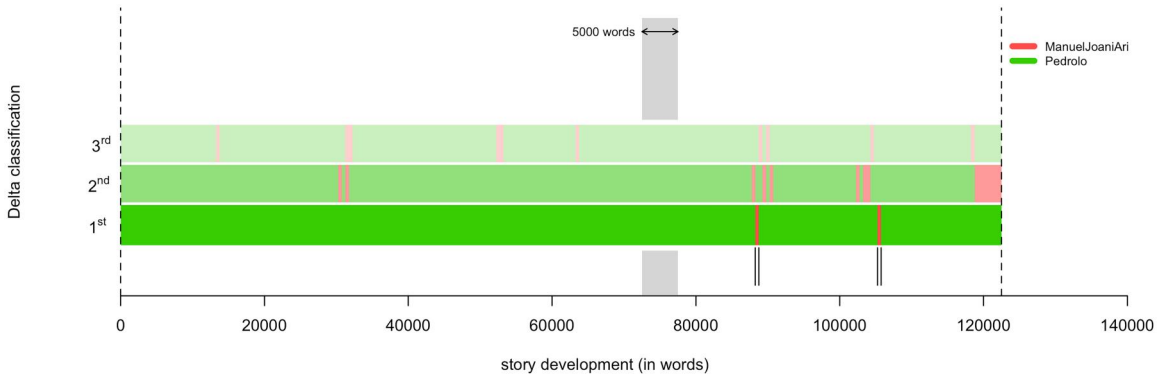


Figure 5. *Els quaderns d'en Marc* trilogy assessed using Eder's rolling stylometry (Delta, 100 MFW).

pair of novels groups together according to its respective author.

When we apply the same settings to a corpus containing the anonymous texts and the works of the four identified potential candidates, the results, as shown in Fig. 2, indicate that the anonymous text clusters with Pedrolo's novels.

To increase the reliability of the authorship attribution exercise, we follow the protocol proposed by Juola (2015). The next step involves introducing an *ad hoc* distractor corpus into the analysis. As mentioned earlier, we compiled a set of 18 works written by Catalan authors active at the time of *Els quaderns*' production but uninvolved in the authorship debate. Applying Eder's

Downloaded from academic.oup.com/dsh/article/41/1/320/8374195 by University College Cork user on 22 October 2026



Figure 6. *Els quaderns d'en Marc* trilogy assessed using Eder's rolling stylometry (SVM, 100 MFW).

Imposters (GI) method (2018), we conduct a second test of the disputed authorship. Following Eder's recommendations, we load this corpus into R *stylo* and draw from a frequency list of the 2,000 most frequent words (MFW). We run the method using three different measurements, Classic Delta, Cosine Delta (Würzburg), and Ruzicka (Kestemont *et al.* 2016). The imposters function produces the following results observed in Table 1.

All candidates except for Manuel de Pedrolo (Ped) and Manel Joan i Arinyó (MA) return very low coefficients across all three distance measures, excluding them from any possible relationship with *Els quaderns*. Even Arinyó's highest score (0.13) is insufficient to suggest a strong link, especially when compared to Pedrolo's consistent score of 1.0 across all measures. This points to a strong resemblance between Pedrolo's authorial fingerprint and that of the anonymous author of the trilogy.

The third stylometric test in our inquiry therefore focuses on these two authors: Manuel de Pedrolo—already the strongest candidate on circumstantial grounds, as argued in previous sections—and Manuel Joan i Arinyó, the only writer showing even a slight resemblance to the *Els Quaderns*' authorial fingerprint. It is also relevant to recall that Joan i Arinyó authored two later sequels to the erotic saga and wrote another partially anonymous work, *Stress* (1985), for the same collection in which *Els quaderns* appeared, as it can be observed in Figs. 3 and 4.[15]

Employing Rolling Stylometry (Eder 2016) with Burrow's Delta, and using the 100 most frequent words (MFW) of each text sample, we explore the possibility that both Pedrolo and Joan i Arinyó collaborated in the writing of the trilogy. One of the distinguishing features of Eder's Rolling Stylometry is its segmentation of the analysed texts; the method attempts to identify which fragments of a text—measured in optimal segments of 5,000 words—may have been written by each potential collaborator. While originally designed to investigate collaboration, the method has also been commonly applied to authorship attribution inquiries, such as Emily Brontë and *Wuthering Heights* (McCarthy and O'Sullivan 2021), Marston and Thomas Nashe and the *Parnassus Plays* (Ilsemann 2018) and Sor Juana Inés de la Cruz and her baroque play *La segunda Celestina* (Hernández-Lorenzo and Byszuk 2023).

Our initial analysis, using the Classic Delta distance, indicates that Pedrolo's authorial fingerprint predominates over Joan i Arinyó's (Fig. 5). However, as Eder notes (2016), the SVM classifier is often a more suitable distance measure for this type of authorship attribution exercise. Figure 6 likewise demonstrates Pedrolo's dominance.

Across the three different stylometric approaches employed in this study, a consistent pattern emerges: Manuel de Pedrolo's authorial fingerprint aligns strongly with that of the anonymous trilogy. Interestingly, when repeating the initial cluster analysis (Cosine delta, 2000 MFW) but this time including the fourth title in the erotic saga, *Els quaderns de la dona d'en Marc* (1985), we obtain some telling results illustrated by Fig. 7.

This fourth instalment of the saga clusters with Joan i Arinyó's works, suggesting that he may have been involved in its production. Following the same procedure as in the previous case, we will now analyse *Els quaderns de la dona d'en Marc* using Rolling Stylometry. While the Delta distance analysis indicates that Joan i Arinyó's authorial fingerprint is closer to that of the anonymous text, the SVM classifier reveals an almost 50/50 distribution, pointing to the possibility of collaborative authorship between Pedrolo and Joan i Arinyó, as shown in Figs. 8 and 9 below.

7. Conclusions

After completing this set of tests with a dataset accurately representing the circle of writers active around this anonymous work, we can confirm what a number of Catalan literary scholars of Pedrolo and twentieth-century Catalan popular literature had already

Downloaded from academic.oup.com/dsh/article/41/1/320/8374195 by University College Cork user on 22 October 2026

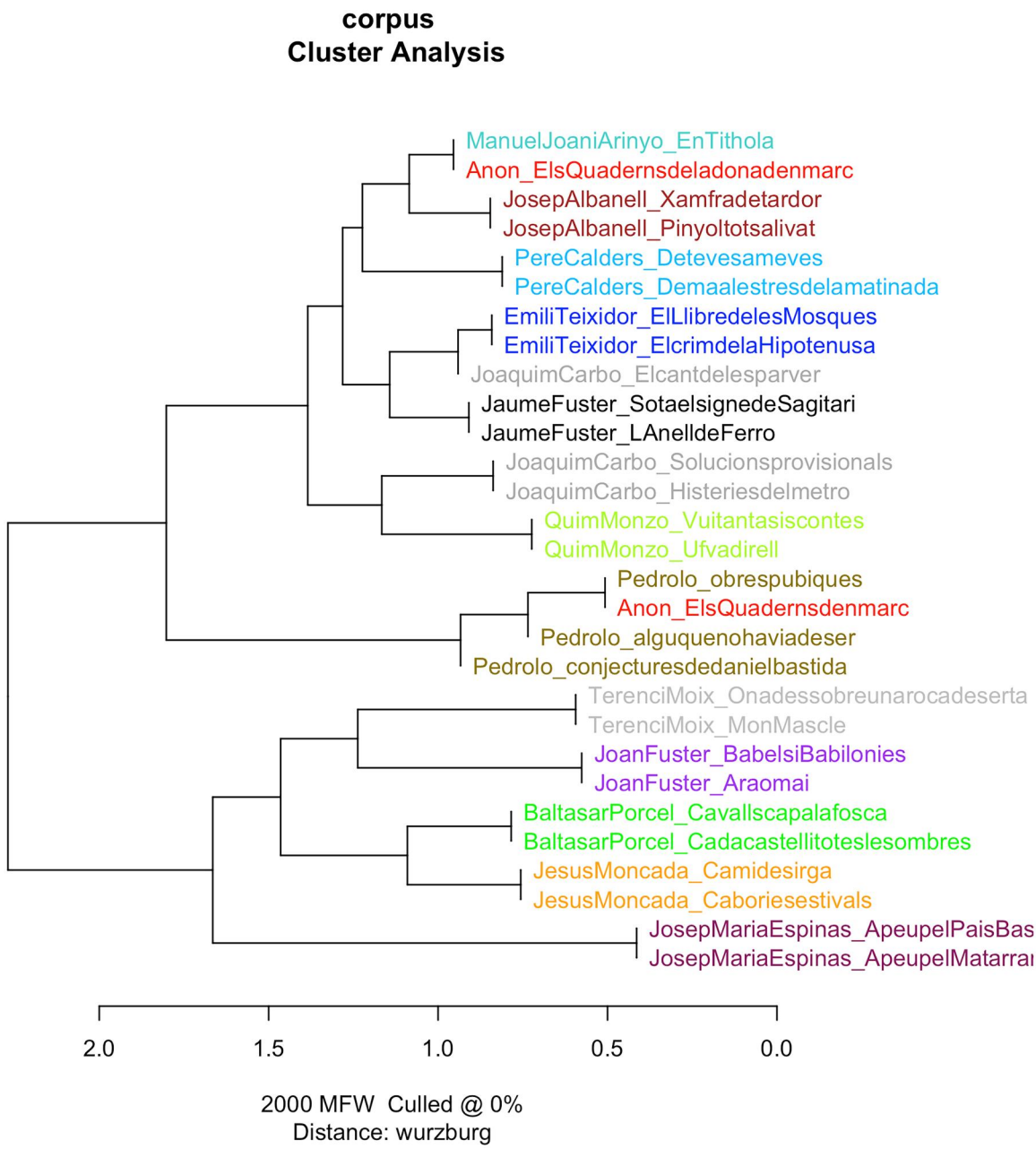


Figure 7. Cluster analysis of the full corpus of potential candidates, *Els quaderns d'en Marc* trilogy and *Els quaderns de la dona d'en Marc* (Cosine Delta, 2000 MFW).

suspected: Manuel de Pedrolo is the most likely author of volumes I, II, and III of *Els quaderns de d'en Marc*. As briefly mentioned in the introduction, this represents a significant addition to Pedrolo's literary legacy: to his pioneering contributions to Catalan crime fiction and science fiction, we may now add his important role in the development of Catalan erotic fiction.

The analysis, however, introduces a new complication. Our results show a strong resemblance between the authorial fingerprint of *Els quaderns de la dona*

Downloaded from academic.oup.com/dsh/article/41/1/320/8374195 by University College Cork user on 22 October 2026

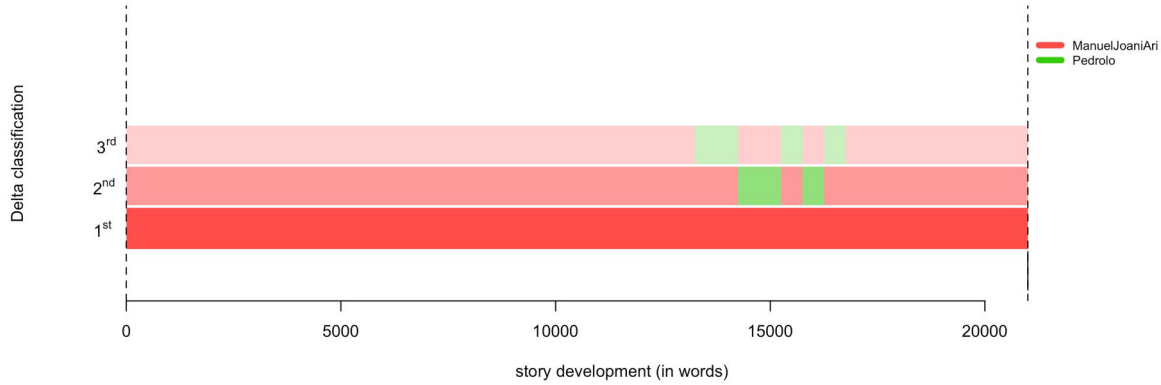


Figure 8. *Els quaderns de la dona d'en Marc* assessed using Eder's Rolling Stylometry (Delta, 100 MFW).

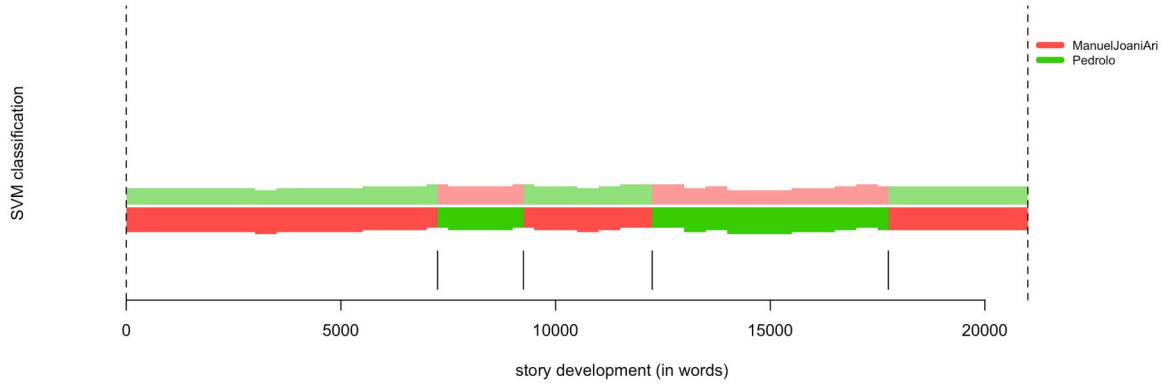


Figure 9. *Els quaderns de la dona d'en Marc* assessed using Eder's Rolling Stylometry (SVM, 100 MFW).

Table 1. Distance measurements across imposters corpus.

Distance	BP	ET	JaF	JM	JoF	JC	JA	JE	MA	Ped	PC	QM	TM
Classic delta	0	0	0	0	0	0	0.05	0	0.13	1	0.03	0.01	0
Cosine delta	0	0.01	0	0	0	0.01	0.05	0	0.12	1	0	0.01	0
Ruzicka	0	0	0	0	0	0.01	0.07	0	0.05	1	0.04	0.03	0

Abbreviations on the table header correspond to the following authors: BP, Baltasar Porcel; ET, Emili Teixidor; JaF, Jaume Fuster; JM, Jesus Moncada; JoF, Joan Fuster; JC, Joaquim Carbó; JA, Josep Albanell; JE, Josep Maria Espinàs; MA, Manuel Joan i Arinyó; Ped, Manuel de Pedrolo; PC, Pere Calders; QM, Quim Monzó; and TM, Terenci Moix.

d'en Marc and that of Manuel Joan i Arinyó. As a writer of later continuations of this saga, as well as another anonymous work in the same collection, the circumstantial evidence aligns with the stylometric findings. It is possible that Joan i Arinyó was the sole author of *Els quaderns de la dona d'en Marc*; alternatively, the near 50/50 stylometric split suggests the possibility of collaboration between him and Pedrolo. Such a collaboration would be highly unusual, as Pedrolo was not known to co-author any other works during his career.

If we return to the introductory note in *Els quaderns d'en Marc* trilogy—now attributed to Pedrolo—where a fictional origin story describes a partnership between an established writer and a younger, aspiring one, we may speculate that this metafictional devide reflects a real dynamic between Pedrolo and Joan i Arinyó. This, however, remains speculation and could form the basis

Downloaded from academic.oup.com/dsh/article/41/1/320/8374195 by University College Cork user on 22 October 2026

for a separate study. For now, we hope that the present research reaches Joan i Arinyó himself, who may be the only one in a position to confirm the true story behind the authorship of the fourth book in the saga.

Author contributions

Pedro Nilsson-Fernàndez (Conceptualization, Data curation, Investigation, Methodology, Resources, Visualization, Writing—original draft, Writing—review & editing)

Data availability

The data underlying this article, mainly the corpus of authors investigated for authorship attribution, cannot be shared publicly due to the works being under copyright. The data will be shared on reasonable request to the corresponding author.

Notes

1. Manuel de Pedrolo will forever be embedded in the Catalan cultural imagination for being the author of the 1974 sci-fi novel *Mecanoscrit del segon origen*, an all-time bestseller in Catalan literature. In addition to this, his crime fiction works achieved considerable success; Catalan noir critics such as Martín Escribà (2018) consider his role as director of the *noir* collection *La cua de palla* to have been instrumental in the development of this genre during the post-war years.
2. As this article will later show in detail, much speculation revolves around Pedrolo's possible authorship of *Els quaderns d'en Marc*. Enric Borràs and Francesc Llaverias claim that Pedrolo wrote the work, while Jordi Arbonès, Pep Albanell, and Joaquim Carbó believe he did not. Other voices, such as Iolanda Pàmies, Joan Josep Isern, and Bel Zaballa, prefer to be more cautious and offer more nuanced opinions on the subject.
3. From now on, *Els quaderns* will be used to refer to volumes I, II, and III of the trilogy. Individual volumes will be specifically named throughout the text when needed.
4. See Mosteller, F. and Wallace, D. L. (1964) *Applied Bayesian and Classical Inference: The Case of the Federalist Papers*. Reading, MA: Addison-Wesley.
5. See Thisted, R. and Efron, B. (1987) 'Did Shakespeare Write a Newly Discovered Poem?', *Biometrika*, 74: 445–55.
6. See Burrows, J. F. (1992) 'Not Unless You Ask Nicely: The Interpretative Nexus Between Analysis and Information', *Literary and Linguistic Computing*, 7: 91–109.
7. See Eder, M., Rybicki, J., and Kestemont, M. (2016) 'Stylometry with R: a package for computational text analysis', *R Journal*, 8: 107–121.
8. By the 1968 date mentioned in this introduction, Manuel de Pedrolo had published a total of twenty-five works. While Pedrolo's corpus would grow to include another hundred titles, twenty-five published works written in the Catalan language is quite an accomplishment in the context of his time. Literature written in Catalan during the late twentieth century was not only in a position of linguistic subalternity with respect to works written in Spanish, but also a political one. Franco's dictatorship (1939–76) imposed the use of the Spanish language as a proposed means of unifying the country. While Francoist censorship affected any form of media regardless of the language in which it was produced, works written in any language other than Spanish were subjected to stricter control measures by the censors.
9. All translations into English from *Els quaderns d'en Marc* and any other references in Catalan are by the author of this piece.
10. *Mecanoscrit del segon origen* (1974) and *Aquesta matinada o potser per sempre* (1976) by Manuel de Pedrolo, are two of the many examples of the rewritten found manuscript trope in his oeuvre.
11. Jordi Arbonès maintained an extended epistolary correspondence with Manuel de Pedrolo from Argentina (where Arbonès resided from 1956 until his death in 2001) and became one of the closest interlocutors to Pedrolo. They shared a firm commitment to Catalan literature, and both were translators of American and English works to Catalan. Arbonès also dedicated part of his literary efforts to promoting Pedrolo's oeuvre, culminating in the extensive study *Pedrolo Perillòs?* (1970).
12. Manuel de Pedrolo experimented with religious themes and tropes in many of his novels. *Mecanoscrit del segon origen* (1974) has a biblical structure and revisits the Adam and Eve trope, while other works such as *Procés de contradicció suficient* (1975/1976), *Crucifeminació* (1981/1986), and *Múltiples notícies de l'Edén* (1984/85) continue this engagement with religious themes.
13. In the context of authorship verification, an 'imposter' is defined as a text known to the researcher not to have been written by the potential author under investigation; imposter corpora are essential when applying the General Imposters framework (GI), introduced by Koppel and Winter (2014).
14. Catalan literature in the late twentieth century continued to bear the effects of nearly four decades of dictatorship, during which languages such as Catalan, Basque, and Galician were heavily censored and, at times, explicitly prohibited by the regime.
15. In addition to being the author of two sequels to this erotic saga, Joan i Arinyó also wrote *Stress* for the collection. This work was first advertised as his own creation in the opening pages of *Els Quaderns d'en Marc III* (1985) but was later presented as anonymous in *Els quaderns de la dona d'en Marc* (1985).

References

Anon. (1984a) *Els Quaderns D'en Marc (I)*. Barcelona: El Llamp.

Anon. (1984b) *Els Quaderns D'en Marc (II)*. Barcelona: El Llamp.

Anon. (1985a) *Els Quaderns D'en Marc (III)*. Barcelona: El Llamp.

Anon. (1985b) *Els Quaderns De la Dona D'en Marc*. Barcelona: El Llamp.

Bacardí, M. (2014) *Epistolari Jordi Arbonès and Joaquim Carbó*. Barcelona: Punctum.

Borràs, E. (2005) *Memòria De Manuel De Pedrolo—Mail Obert—VilaWeb*. VilaWeb.cat.

Carbó, J. (2005) 'La Novel·la De Jordi Arbonès', *Quaderns. Revista De Traducció*, **12**: 11–24.

Eder, M. (2016) 'Rolling Stylometry', *Digital Scholarship in the Humanities*, **31**: 457–69.

Eder, M. (2017) 'Visualization in Stylometry: Cluster Analysis Using Networks', *Digital Scholarship in the Humanities*, **32**: 50–64.

Eder, M., Rybicki, J., and Kestemont, M. (2016) 'Stylometry with R: A Package for Computational Text Analysis', *The R Journal*, **8**: 107–121.

Efron, B., and Thisted, R. (1976) 'Estimating the Number of Unseen Species: How Many Words Did Shakespeare Know?', *Biometrika*, **63**: 435–47.

Evert, S. *et al.* (2017) 'Understanding and Explaining Delta Measures for Authorship Attribution', *Digital Scholarship in the Humanities*, **32**: ii4–ii16.

Hernández-Lorenzo, L., and Byszuk, J. (2023) 'Challenging Stylometry: The Authorship of the Baroque Play La Segunda

Downloaded from academic.oup.com/dsh/article/41/1/320/8374195 by University College Cork user on 22 October 2026

Celestina', *Digital Scholarship in the Humanities*, **38**: 544–58.
Ilsemann, H. (2018) 'Stylometry Approaching Parnassus', *Digital Scholarship in the Humanities*, **33**: 548–556.
Isern, J. J. (2001) 'Un Almogàver Sota Els Llençols', *Avui*, **12**.
Joan i Arinyó, M. (1985) *Stress*. Barcelona: El Llamp.
Kestemont, M. *et al.* (2016), 'Authorship Verification with the Ruzicka Metric', in *Digital Humanities 2016: Conference Abstracts*, 246–49. Kraków: Jagiellonian University and Pedagogical University.
Juola, P. (2015) 'The Rowling Case: A Proposed Standard Analytic Protocol for Authorship Questions', *Digital Scholarship in the Humanities*, **30**: fqv040.
Llaverias, C. (2016) *Els quaderns d'en Marc; Catalunya serà eròtica o no serà* (Barcelona). https://llibressenyordolent.blogspot.com/2016/06/els-quaderns-den-marc.html, accessed November 2024.
Martín Escribà, À. (2018) *La Novel·la Negra i Policíaca en Llengua Catalana, LletrA—La Literatura Catalana a Internet*. https://lletra.uoc.edu/ca/tema/novel-la-negra-i-policiaca/detall, accessed 26 June 2024.
McCarthy, R., and O'Sullivan, J. (2021) 'Who Wrote Wuthering Heights?', *Digital Scholarship in the Humanities*, **36**: 383–391.
Miller, H. (1956) *A Devil in Paradise*. New York: Signet.
Miller, H. (1975) *My Life and Times*. New York: Playboy Press.
Miller, H. (1983) *Under the Roofs of Paris*. New York: Grove Press.
Miller, H. (1992) *Sexus*, Edicions 62, Barcelona.
Miquel. (2015) *Provisional: L'autoria dels Anònims o Catalunya Serà Eròtica o No Serà!* https://provisionals.blogspot.com/2015/05/lautoria-dels-anonims.html, accessed 20 June 2024.
Nilsson-Fernàndez, P. (2019) *Mapping the Works of Manuel de Pedrolo in Relation to the Post-Civil War Catalan Landscape* **[Doctoral Thesis]**, University College Cork, print. Available at: https://hdl.handle.net/10468/8959. Accessed June 2024.
Ochab, J. K. *et al.* (2019) 'Identifying Similarities in Text Analysis: Hierarchical Clustering (Linkage) versus Network Clustering (Community Detection)', In *Digital Humanities 2019: Book of Abstracts*. ADHO: Utrecht.
Pagano, A. *et al.* (2024) 'Authorship Attribution in Translated Texts: A Stylometric Approach to Translator Style', in P. Gamallo ***et al.*** (ed.) *Proceedings of the 16th International Conference on Computational Processing of Portuguese—Vol.* **2**, pp. 119–117. Association for Computational Linguistics: Santiago de Compostela.
Pàmies, I. (2010) *Mots Muts Nats: El Secret Més Ben Guardat de la Literatura Catalana … ?*. https://motsmutsnats.blogspot.com/2010/01/el-secret-mes-ben-guardat-de-la.html, accessed 25 June 2024.
Pijuan Vallverdú, A. (2016) *Manuel de Pedrolo, traductor de poesia*, Doctoral Thesis, Universitat de Barcelona.
Rybicki, J. (2013) 'Stylometryczna Niewidzialność Tłumacza', in *Przekładaniec: A Journal of Literary Translation*, **27**. Wydawnictwo Uniwersytetu Jagiellońskiego.
Zipf, G. K. (1932) *Selected Studies of the Principle of Relative Frequency in Language*. Cambridge, Mass: Harvard University Press.

Downloaded from academic.oup.com/dsh/article/41/1/320/8374195 by University College Cork user on 22 October 2026

© The Author(s) 2025. Published by Oxford University Press on behalf of EADH.
This is an Open Access article distributed under the terms of the Creative Commons Attribution-NonCommercial License (https://creativecommons.org/licenses/by-nc/4.0/), which permits non-commercial re-use, distribution, and reproduction in any medium, provided the original work is properly cited. For commercial re-use, please contact reprints@oup.com for reprints and translation rights for reprints. All other permissions can be obtained through our RightsLink service via the Permissions link on the article page on our site—for further information please contact journals.permissions@oup.com.
Digital Scholarship in the Humanities, 2026, 41, 320–332
https://doi.org/10.1093/llc/fqaf096
Short Paper