Reconstructing Readers' Tastes through Topic Modeling of Private Poetry Anthologies

Álvaro Ceballos Viro

Université de Liège

Abstract: Throughout the nineteenth and twentieth centuries, many lay readers commonly transcribed poems that were meaningful to them by hand into notebooks. I call these cultural artifacts 'private poetry anthologies'. The heterogeneity of the anthologized texts in these notebooks makes it difficult to understand what uses or aesthetic preferences they responded to. I argue that topic modeling can be a useful digital tool for identifying thematic patterns specific to these materials. This article presents the results of analyzing eight such notebooks in Spanish language, using a control corpus of 54 poetic works from the same period for comparison. The results reveal that the private anthologies tend to emphasize sentimental themes, and that overall they are not outliers but rather tend to be close to the thematic mean of the poetry of their time.
Keywords: Spanish poetry, lay readers, topic modeling

1. Introduction

The challenges of studying the empirical reception of literature increase as we look further back in time. Going back just a few generations, there is no way to find out what readers in the past did with literary texts. We cannot conduct surveys or biometric tests on

readers from the Belle Époque, which is one of the periods I have studied most closely. There are a few rather indirect and often uncertain ways of reconstructing real everyday reading experiences, like marginalia (Sherman 2008), personal diaries or memoirs,[1] correspondence (Lyon-Caen 2006; Rohrbach 2017), and, occasionally, court records — I am thinking, of course, of Carlo Ginzburg's famous Menocchio and the way he dealt with the handful of books he managed to get hold of (1992), but also of Antonio Castillo's creative use of the Inquisition archives (2016). In this context, one documentary source that has been almost completely overlooked is what I call 'private anthologies of poetry'.

Throughout the nineteenth and twentieth centuries, many lay readers commonly transcribed their favorite poems, or poems that were emotionally or epistemologically meaningful to them, by hand into notebooks. This cultural practice is linked to the popularization of the notebook as a technological device, initially put together from loose sheets of paper by the readers themselves, later purchased from a commercial company. It is not uncommon for desk products intended for other uses, such as diaries or bookkeeping notebooks, to be repurposed as a material support for private poetry anthologies.

The well-known narratologist Gérard Genette recounts how his mother had a notebook in which she had transcribed a series of poems over the years, and how she made Genette himself read them aloud. Some of these poems had enjoyed enormous popularity in France in the 1900s; others rhymed with intimate, traumatic

---

[1] Like the 30,000 citations that attest to historical reading experiences documented by The UK Reading Experience Database, www.open.ac.uk/Arts/reading/UK/index.php.

experiences — the death of a child — that the narratologist's mother had gone through (2006, 50-51).

Private poetry anthologies must have served a wide array of non-exclusive purposes such as preserving a fading repertoire of memorized works, conserving and appropriating hard-to-find texts, circumventing censorship, displaying one's cultural capital, etc. In any case, the poems *spoke* to their owners, and the owners *spoke* through them — even if only to themselves.
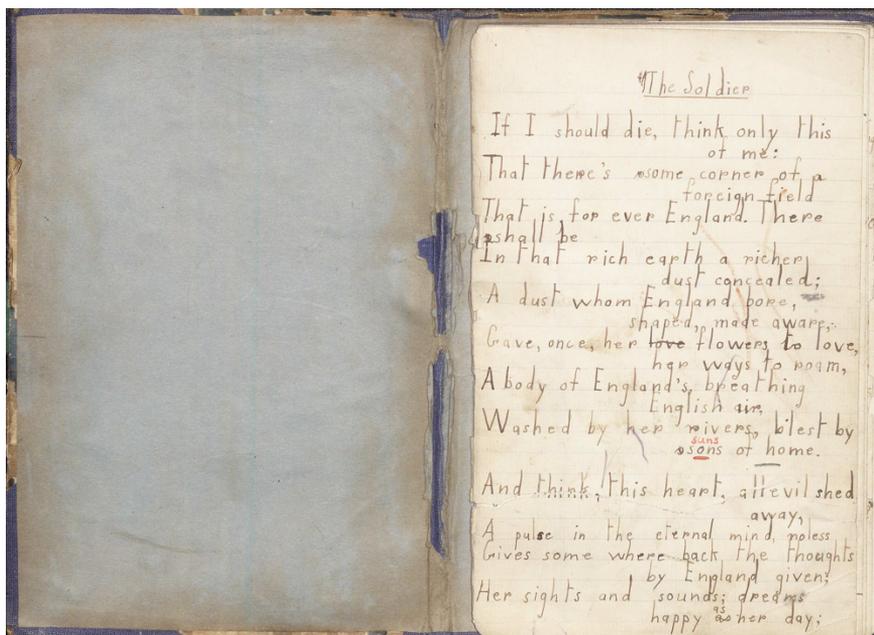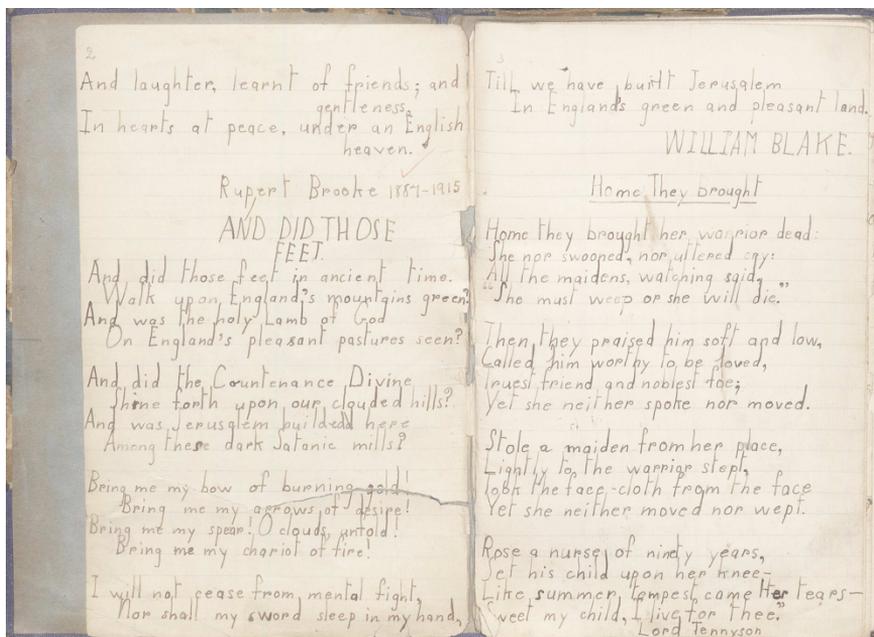
Illustrations 1 and 2:
An English private anthology from around 1948 (excerpt).

Interestingly, the identities of the transcribers are often prominently displayed in these private anthologies — on the title page, for example, or even at the bottom of each transcribed poem.

This makes the metaphor of the text *appropriation* through reading feel more tangible, as if the voice of the poem merged with the voice of the person copying it and naturally fitted into a mental experience that felt personal and intimate. One of the notebooks in my collection (Illustrations 1 and 2) contains, after many poems by Kipling, Yeats, Wordsworth, Browning, and, above all, John Masefield, a note dated 1948 that reads: "Gladys O. is my / name England is my nation / Heven [*sic*] is my dwelling place / Christ is my salvation / When I am dead and my / bones are roten [*sic*] I'll leave / you this note for you to read /

when I am quite forgotten [*sic*]." Despite their naivety and — no doubt — the playful and ritualistic nature of these rhymes, they can be seen as a hint of how copying poems, mimicking an authorial pose, could satisfy a need for transcendence, ensuring that death does not completely erase one's existence.

These collections offer a striking contrast to the established literary canon and linear narratives of literary history, uniquely blending high and low literature, incorporating works from different centuries, countries, and sometimes even languages. Typically, these transcriptions were not exact copies of an edited version, but often bore traces of oral transmission and selective memory. It is not uncommon to find crossed-out words, or jottings that correct memory lapses, or blank spaces intended for later completion of half-forgotten poems. On other occasions, however, it can be proved that a poem has been copied directly from printed material, and there are even cases where a poem torn from a magazine or a clipping from broadside ballads has been glued directly onto the paper.

What makes these private anthologies extraordinary is that they represent collections curated not according to historiographical or thematic criteria (as seen in commercial or critical anthologies), but rather based on a form of personal *attachment* to specific poems — in the multilayered sense Rita Felski gives to this term (2020). They share similarities with school practices (Escolano 2000, Eddy 2018), intimate writing, collectibles, German or Dutch *Poesiealben* (Walter 2019), writers' reading notebooks (Minzetanu 2016), and other everyday writing practices. Despite their cultural significance, these notebooks have received little scholarly attention and have never been studied systematically.

These artifacts, devoid of commentary, are as captivating as they are opaque; it is challenging to make them speak about the everyday

uses of poetry, as such uses are revealed almost exclusively through the selection of texts.

Could it be that, beyond sharing generic, material, and formal characteristics, they also share semantic features? I hypothesize, indeed, that such collections feature specific semantic networks. I decided to subject some of these private anthologies to stylometric analysis, in order to identify patterns not immediately visible, hoping they could tell us something about the usual preferences of lay readers.

2. Experimental and control corpus

As a researcher of Romance languages, I selected eight notebooks written in Spanish between 1850 and 1940.[2] All of them have been digitized and transcribed on the Transkribus platform.[3] I then added a control corpus consisting of 54 original poetry works published in Spanish between 1826 and 1939. Many of them come from digital libraries such as the Project Gutenberg or Wikisource, but I have occasionally transcribed digital versions of printed copies with the help of PDF2Go (a platform that offers AI-assisted OCR and, unlike others, does not store documents for later use). All titles in the control corpus are in the public domain. I tried to ensure that this control corpus was representative of the different poetic genres that circulated during the period. Therefore, it not only contains titles by the most widely read lyric poets within educated circles (such as Rubén Darío or Ramón de Campoamor), but also satirical poetry (a selection of the vast body of work by Luis de Tapia), light verse

---

[2] The scanned originals are available in the following repository: https://doi.org/10.17605/OSF.IO/8NJXQ (reference numbers 1, 3, 4, 6, 8, 9, 10, and 12).

[3] I would like to thank Adrià Aguacil Portillo and Gerardo O. Dávila Llamas for their collaboration.

(Antonio Palomero, Juan Pérez Zúñiga), broadsheet ballads, regionalist and religious poetry, both Spanish and Latin American. Five of the authors represented were women (i.e., 9.25%) — consider as a reference point that, in 1922, only 32 of 407 published authors, old or new, Spanish or foreign, were women (i.e., 7.8%).[4] The dates identifying the texts in the corpus are approximate: in some cases (for example, the poems by Rosalía de Castro, but also those by Pérez Zúñiga and Antonio Palomero), I resorted to volumes of complete or selected works that were originally written several years before publication. The control corpus totals 6.9 MB of plain text; the experimental corpus (the private anthologies) totals 0.496 MB (6.71% of the total corpus).

Depending on their source, the resulting files did not always maintain the correct verse breaks or line divisions, but this is not relevant to semantic analysis. All texts have been manually curated; I removed headers, page numbers, Roman or Arabic numerals, and editorial peritexts. I smoothed the most frequent diachronic differences (*fué, dió, pié, á, ó, é, tí, fé, -cion, -sion*). Even so, the result is not entirely systematic, and some texts still present dialectal marks that I did not want to remove.

I have relied on Claude Sonnet 4 for the orthographic normalization of four of the titles (those by Robustiana Armiño, the Marquis of Cerralbo, José Selgas, and Antonio Palomero). One of the private anthologies in the experimental corpus, written by someone on the threshold of literacy and riddled with misspellings, abbreviations, and transcription errors, has also been normalized using the same AI model; this has resulted in spelling and morphology comparable to

---

[4] Self-collected data related to literary publications included in the bulletin *Bibliografía Española*. The women in my control corpus are Delmira Agustini, Robustiana Armiño, Rosalía de Castro, Antonia Díaz de Lamarque, and Ángela Grassi.

the other documents, at the expense of creative variants that, while not devoid of interest, often rendered the text unintelligible and prevented any comparison. For example: "Murmurando un manzo río / baño el prado con sociego [...] / tu in cilencio anto mío / mi afligido pecho bañas" was transformed into this more coherent and probable passage: "Murmurando un manso río / baña el prado con sosiego [...] tú en silencio santo mío / mi afligido pecho bañas". Delegating normalization to a probabilistic model produces a text that is both plausible and close to the corrupted version, avoiding the need to make case-by-case decisions that may lack consistency.

I chose to lemmatize with Stanza, an NLP library developed by Stanford, which I believe outperforms the more popular SpaCy. Despite occasionally maintaining gender agreement in Spanish adjectives, Stanza is more systematic in matching words to their lemmas, while the SpaCy library for Spanish (es_core_news_lg) is inconsistent with some verb forms, especially participles and gerunds. Although both libraries occasionally generate incorrect lemmas, SpaCy is more prone to doing so — producing pseudowords such as 'enjambr,' or 'alegrísir'.

I applied a list of 238 stopwords, whose suitability for this already lemmatized corpus had been previously verified in Voyant. Given that the items in the corpus do not have comparable dimensions, I have divided them into 3,326 chunks with an average of 198.9 words — 200 words is an empirical sweet spot for LDA (*cf*. Cooper *et al.* 2021). This should make the analysis more robust and lead to more homogeneous topics.

3. Topic modeling
Topic modeling refers to algorithms that identify which words are likely to co-occur in a textual corpus. The best known of these

algorithms is Latent Dirichlet Allocation (LDA). The result of this mathematical operation is a series of topics, i.e., lists of words that tend to interact in the corpus, many of which have a clear semantic or situation-based connection that we can think of as a proxy for a 'theme' in that corpus — 'theme' meaning the choice an empirical reader makes from one of the objectifiable isotopies in a text or series of texts.

Using the Mallet wrapper designed by Maria Antoniak,[5] I trained an LDA model with different numbers of topics, ranging from 10 to 30. The number of topics with the highest coherence score — calculated using the Gensim library — turned out to be 30.

number of topics	Coherence score
10	0.4683
15	0.4694
20	0.4827
25	0.4724
30	0.4899

Of those 30 topics, 18 were manually validated, as they presented strong semantic and contextual coherence (topics 0, 1, 3, 4, 5, 7, 9, 13, 14, 15, 18, 20, 21, 22, 23, 24, 25, and 28)[6]:

Nº	WEIGHT	TOP WORDS	TOP WORDS (ENGLISH)	LABEL
0	0.27738	alma noche sombra tarde viejo triste blanco soñar luna sueño sol amor decir vida	soul, night, shadow, evening/late, old, sad, white, dream [verb], moon, dream	nighttime melancholy

		hora	[noun], sun, love, say, life, hour	
1	0.12486	santo humilde dulce maría esperanza madre celestial virgen piedad consuelo dios pueblo ofrecer llegar vano	holy, humble, sweet, Mary, hope, mother, heavenly, virgin, mercy, consolation, god, people, offer, arrive, vain	Virgin Mary
2	0.0339	dios mar españa camino buen día señor corazón don hombre mañana soñar nada llevar ayer	god, sea, Spain, road, good, day, lord, heart, Mr./talent, man, tomorrow, dream, nothing, carry, yesterday	?
3	0.31165	gloria dios mundo alma hombre divino tierra luz sol inmortal cielo nombre santo día eterno	glory/heaven, god, world, soul, man, divine, earth, light, sun, immortal, sky/heaven, name, holy/saint, day, eternal	religious paradise
4	0.02402	libertad pueblo patria cabeza frente medio tirano levantar decir vez pampa pié américa mayo libre	freedom, people, homeland, head, front/forehead, middle, tyrant, raise, say, time [as in 'two times'], pampa, foot [old spelling], America, May, free	American liberation
5	0.09799	hombre sangre toro piedra boca dar hueso perro mano cuerpo corazón hasta carne hambre diente	man, blood, bull, stone, mouth, give, bone, dog, hand, body, heart, until, flesh, hunger, tooth	visceral quality
6	0.04413	ustar usté mucho cuatro cosa señora tomar calle tres mismo cualquiera persona soler buen seis	[?], you/sir, a lot, four, things, madam, take, street, three, same, any, person, usually, good, six	?
7	0.01655	indio gaucho andar mesmo aunque cantar hombre vez hasta cuanto mucho pobre ansi sufrir fierro	indian, gaucho, walks, same [old spelling], even, sing, man, time [as in 'two times'], until, how, as … as, much/a lot, poor, thus/so [old spelling], suffer, iron [old spelling]	*gauchesca* (gaucho literature)
8	0.63999	decir ver dar tanto mucho querer día pasar venir buen hombre así poner llegar poco	say, see, give, so much, much/a lot, want/love, day, pass, come, good, man, thus/so, put, arrive, little	?
9	0.06098	rey granada moro cristiano tal sultana muley torre zoraya decir alhambra mora árabe esclavo alhama	king, Granada, moor, christian, such??, sultan's wife, muley [sultan, Arabic prince], tower, Zoraya [Middle-East	Arab setting

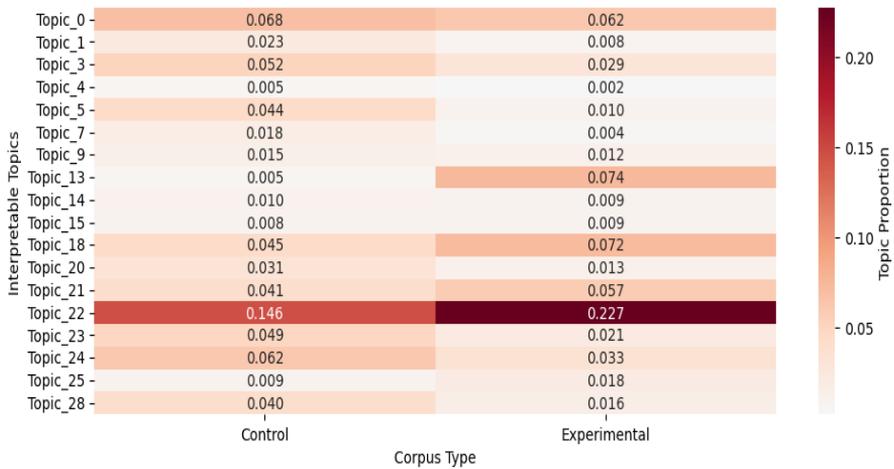|			first name], say, Alhambra, moor woman, Arab/Arabic, slave, Alhama [Arabic place name]	
10	0.03036	marido san puerta muger viejo casa mujer cuarto punto mucho juan prado subir casar medio	husband saint door woman [old spelling], old, house, woman, room/fourth, point, much/a lot, Juan, field, go up, marry, half	?
11	0.27937	sombra ojo noche silencio voz luz puerta paso mano vez pie dejar tanto rostro brazo	shadow, eye, night, silence, voice, light, door, step, hand, time [as in 'two times'], foot leave, so much, face, arm	?
12	0.05914	hoy don lector vivir ayer tal vez nuestro juan gente luis aunque gran gato cierto	today, Mr./talent, reader, live, yesterday, such, time [as in 'two times'], our, Juan, people, Luis, although, big/grand, cat, certain/true	?
13	0.04084	amor querer amar beso mujer mía muero bello pasión amante vivir dulce corazon morir adorar	love [noun], want/love, love [verb], Kiss, woman, mine, I die, beautiful, passion, lover, live, sweet, heart, die, adore	hyperbolic passion
14	0.05605	nuestro españa raza nota filipino bandera rizal flor ideal nuevo español madre idea vivir américa	our, Spain, race, note, Filipino, flag, Rizal [Filipino national hero], flower, ideal, new, Spanish, mother, idea, live, America	colonial Spain
15	0.03728	rey princesa vuestro oro palacio noble oso toledo paje acmed alcázar señor caballero conde castillo	king, princess, your, gold, palace, noble, bear, Toledo, pageboy, Acmed, alcazar [palace], lord, knight, count, castle	courtly atmosphere
16	0.07239	don decir casa dios gran mucho padre dama caballero señor juan vuestro quedar pedir muerte	Mr./talent, say, house, god, great/big, much, father, lady, gentleman/knight, lord, Juan, your [plural], stay, ask, death	?
17	0.00521	andar andir probe frío sendico páece too allí cuatro tós jacer juar malo juan huaco	walk, walk [?], poor [vulgar], cold, trustee [vulgar], seems [vulgar], everything [vulgar], there, four, everyone [vulgar], do [vulgar], [?], bad, Juan, [?]	?

1 8	0.24411	flor amor rosa tanto ojo hermoso bello cantar blanco labio mujer sol niña dar belleza	flower, love, rose, so much, eye, beautiful, pretty, sing, white, lip, woman, sun, girl, give, beauty	female beauty
1 9	0.01167	clara calixto der cuba soldado industria arte galar peso árbitro regocijo vario cama fatiga ondi	Clara, Calixto, [?], Cuba, soldier, industry, art, [?], weight, arbitrator, rejoice, various, bed, fatigue, [?]	?
2 0	0.10878	tierra campo camino viejo sol agua blanco monte río verde tarde alvargonzález casa rama hijo	land, field, path, old, sun, water, white, mountain, river, green, afternoon, Alvargonzález, house, branch, son	rural setting
2 1	0.25094	sangre patria españa pueblo guerra nuestro gloria frente noble vuestro fuerte mano espada honor muerte	blood, homeland, Spain, people, war, our, glory, front/forehead, noble, your [plural], strong, hand, sword, honor, death	patriotic war
2 2	0.71982	amor alma vida triste corazón dolor tanto ver sólo siempre vez ojo mundo llorar querer	love, soul, life, sad, heart, pain, so much, see, only, always, time [like in 'two times'], eye, world, cry/sob, want/love	heartbreak
2 3	0.25394	mar viento cielo sol nube rayo ola inmenso monte tierra onda mundo nave luz mare	sea, wind, sky, sun, cloud, lightning, wave, immense, mountain, earth, wave, world, ship, light, sea?	seaside atmosphere
2 4	0.31678	flor dulce ver ave luz cielo allí sol agua fuente ala mil aura frente hermoso	flower, sweet, see, bird, light, sky, there, sun, water, fountain, wing, thousand, aura, forehead/front, beautiful	*locus amœnus*
2 5	0.05569	don caballo espada caballero acero juan capitán inés diego ancho muerte fiero sol lanza arma	Mr./talent, horse, sword, knight, steel, Juan, captain, Inés, Diego, wide, death, fierce, sun, spear, weapon	cavalryman
2 6	0.08464	tanto vuestro duro virtud mal gente don alto campo vano oro fiera amistad monte envidia	as much, your [plural], hard, virtue, bad, people, Mr./talent, high, field vain gold fierce friendship mountain envy	?
2	0.00808	loro mico per cuyo mes	parrot, monkey, for [catalan],	?

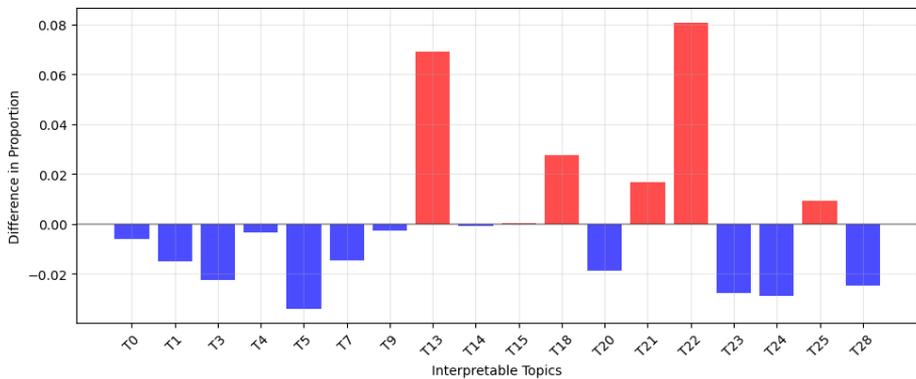| 7 | | moro allá rico amb bella vaig puerto com pobre isola | whose, month, there, rich, with [catalan], beautiful, go [catalan], port, as [catalan], poor, island [italian] | |
| 28 | 0.11229 | rosa oro divino blanco gran vida verso musa luz flor boca alma ensueño mano ala | rose, gold, divine, white, great, life, verse, muse, light, flower, mouth, soul, daydream, hand, wing | romantic lyricism |
| 29 | 0.07415 | luna niño agua sangre amor cantar negro cielo aire amargo caballo venir ojo verde tres | moon, child, water, blood, love, sing, black, sky, air, bitter, horse, come, green, eye, three | ? |

Some of the incomprehensible terms in the table above ('ustar,' 'der,' 'galar') may be corrupted forms resulting from the lemmatization process; it should also be remembered that the corpus also contains diatopic (dialectal) or diastratic variants that might be difficult for an NLP library to identify. I hesitated to keep some of the topics, especially numbers 5 and 28. Topic number 5 is semantically obscure, but many of its words refer to animality and the physical materiality of the body; topic 28 does not belong to a clear semantic field or social script, but it does bring together many positively connoted terms commonly used in the expression of romantic feelings. I discarded topic 29, although anyone familiar with the poetry of Federico García Lorca will recognize several of his most recurrent and prominent symbols in it.

4. Results

The heat map of these 18 topics reveals expected patterns: the prominence of the gaucho topic (7) in the works of Argentine José Hernández; that of topic 20 ('rural setting') in the two works by Antonio Machado; that of the Arabizing topic (9) in one of the volumes of works by José Zorrilla; that of 'romantic lyricism' in the works of

Rubén Darío… and the relative absence of *all of them* in the poems of the satirist Luis de Tapia and in the broadside ballads.

Graph 1: heat map of topic distributions

Five topics are more prominent in the experimental corpus than in the control corpus (Graphs 2 through 4). Each of them appears in different proportions across the private poetry anthologies (Graph 5). In this sense, topic modeling can also be used to identify specific preferences of particular readers, through materials such as those studied here. Notably, topic 22 (which I have titled 'heartbreak') is particularly prominent in all of them, while number 21 ('patriotic war'), although present in others, is predominant in a notebook written in Cuba during the years of the Spanish–American War.

Graph 2: topic distributions, experimental vs. control corpus

Graph 3: topic prominence, experimental (red) vs. control (blue)
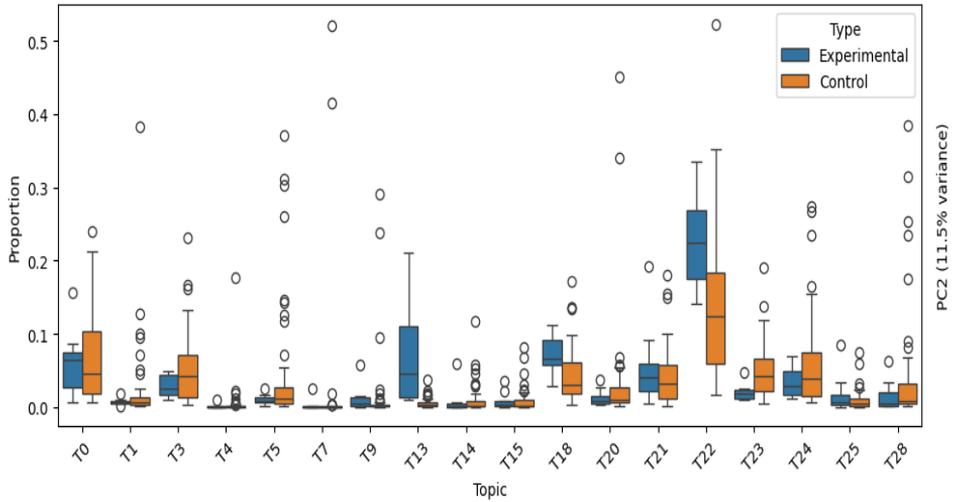
Graph 4: topic distribution variability by document type
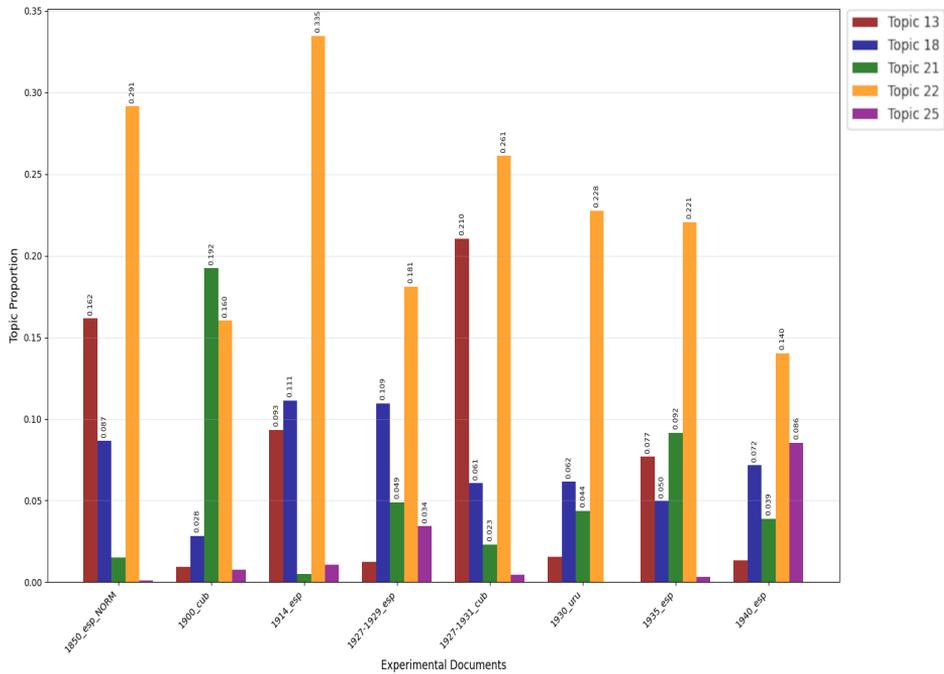
*Graph 5: distribution in experimental corpus of topics 13, 18, 21, 22, and 25*

Only three of these five topics show statistical significance (with a one-tailed p-value < 0.05): topics 22, 13, and 18. By contrast, topics 21 and 25 should be discarded. Note, furthermore, that when applying a conservative correction for multiple comparisons (Holm), only topic 22 maintains robust significance (0.0290). The effect size (Cohen's d) of topic 13 is very large: this topic is notably present in several documents in the experimental corpus and virtually absent from the control corpus. The effect size of topics 18, 22, and 25 is also large.

topic	label	p-value	Cohen's d
13	'hyperbolic passion'	0.0195	2.550
18	'female beauty'	0.0167	0.750
21	'patriotic war'	0.2354	0.392
22	'heartbreak'	0.0058	0.804
25	'cavalryman'	0.1981	0.577

5. Discussion and limitations

I believe that the results are highly interpretable. All three statistically significant topics show a clear semantic relationship: 'hyperbolic passion,' 'female beauty,' and 'heartbreak' are three typical facets of Western love poetry, three dimensions of the same theme. Private poetry anthologies would reveal a perhaps expected preference for this type of content or musing. Indeed, some poems, perhaps even entire notebooks, appear to have been used in a context of seduction. This seems to be suggested by numerous clues in the oldest of the notebooks considered here, decorated with hearts and rich in love poems (Illustration 3), as well as by the gallant poems in the Cuban anthology — such as the one reproduced below, which reads, in

acrostic form, 'dame un beso' ('kiss me') and, according to a handwritten footnote, was sent consecutively to two young ladies (Illustration 4).
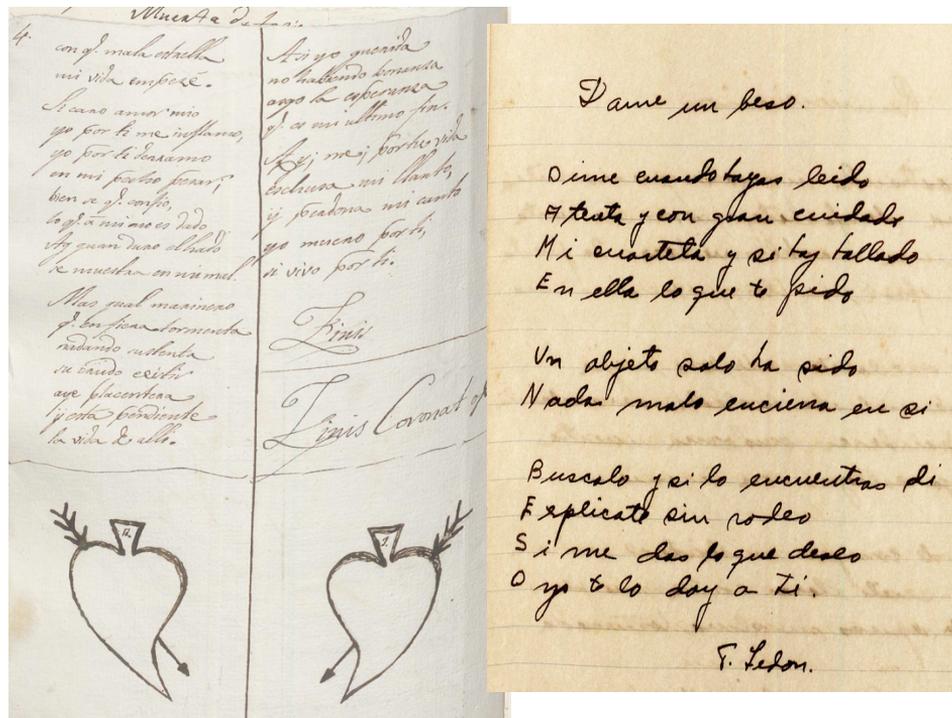
Illustrations 3 (l.) and 4 (r.):
Romantic features in two of the analyzed notebooks

These private anthologies also represent different trends in the poetry of the period, but overall they are not outliers. Principal component analysis (PCA, Graph 6) shows that, although they emphasize sentimental aspects, they tend to be close to the thematic mean values of the control corpus.
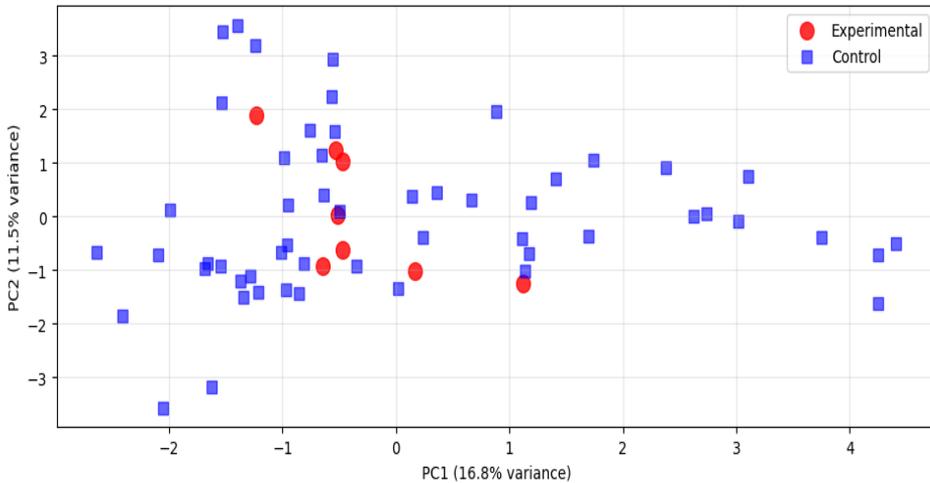
Graph 6: Principal component analysis (PCA): two-dimensional document clustering

These results are in line with a pilot study conducted with Dariah-Topics Explorer on half of the documents (Ceballos, 2025). However, I am well aware of the limitations of the present study. To begin with, the noise introduced by lemmatization shortcomings might be reduced by developing new and improved libraries for Spanish. The size of the corpus influences the robustness of the analysis, and I know that 62 documents, even though many contain hundreds of pages, are not sufficient to make strong claims about the preferences of poetry readers in the so-called 'Silver Age' of Spanish literature (*cf.* Calvo Tello and Rißler-Pipka 2022). Furthermore, I have analyzed private anthologies written in quite different periods and in very different regions (most of them come from Spain, but two were written in Cuba and one in Uruguay): the corpus would need to be considerably expanded to confirm the existence of transhistorical and

transnational similarities. Unfortunately, the scarcity of these archival materials makes it impossible to carry out a more ambitious study in the short term. Still, I believe that the results of this research show the great potential of this methodology when it comes to interpreting such elusive accounts of everyday reading as private anthologies.

Data available on
https://doi.org/10.5281/zenodo.17054856
https://doi.org/10.17605/OSF.IO/8NJXQ

References

Ceballos Viro, A. 2025. "Poetry Lovers of the Past: Topic Modeling of Private Anthologies." Scientific talk in the conference "Impressed by Reading. Measuring the Impact." Leiden, 24–27 February.

Calvo Tello, José, and Nanette Rißler-Pipka. 2022. "Replication Crisis and the (Digital) Humanities: Perspectives from the Spanish Silver Age(s)." In *Towards the Digital Cultural History of the Other Silver Age Spain*, edited by Dolores Romero López and Jeffrey Zamostny. Peter Lang. 35–66.

Castillo Gómez, Antonio. 2016. *Leer y oír leer. Ensayos sobre la lectura en los Siglos de Oro*. Iberoamericana and Vervuert.

Cooper, A. F. *et al*. 2021. "'Tecnologica cosa'. Modeling Storyteller Personalities in Boccaccio's *Decameron*." *Proceedings of LaTeCH-CLfL 2021*, online. https://doi.org/10.18653/v1/2021.latechclfl-1.17

Eddy, Matthew D. 2018. "The Nature of Notebooks: How Enlightenment Schoolchildren Transformed the Tabula Rasa." *Journal of British Studies* 57: 275–307.

Escolano, Agustín. 2000. "La copie de manuscrits comme pratique scolaire d'appropriation et ses usages de l'écriture." In *Copie et modèle: usages, transmission, appropriation de l'écrit*, edited by

Christine Barré-De Miniac. Institut National de Recherche Pédagogique, 41–56.

Felski, Rita. 2020. *Hooked. Art and Attachment*. The University of Chicago Press.

Genette, Gérard. 2006. *Bardadrac*. Seuil.

Ginzburg, Carlo. 1992 (1976). *The Cheese and the Worms. The Cosmos of a Sixteenth-Century Miller*. The Johns Hopkins University Press. Translated by John and Anne Tedeschi.

Lyon-Caen, Judith. 2006. *La Lecture et la vie. Les usages du roman au temps de Balzac*. Tallandier.

Minzetanu, André. 2016. *Carnets de lecture. Généalogie d'une pratique littéraire*. Presses Universitaires de Vincennes.

Rohrbach, Véronique. 2017. *Le Courriel des lecteurs à Georges Simenon. L'ordinaire en partage*. Presses Universitaires de Rennes.

Sherman, William R. 2008. *Used Books. Marking Readers in Renaissance England*. University of Pennsylvania Press.

Walter, Stefan. 2019. *Der Staat und die Werte: Wertwandel in Poesiealben in der DDR und Bundesrepublik 1949–1989*. Springer VS.