Lexicon, Affect, and Meter: Quantitative Insights into the Poetics of Antonio and Manuel Machado

Sunghyun Jang & Hye-Yoon Chung

Abstract: This study employs a data-driven framework to offer a computational re-evaluation of the poetry of Antonio and Manuel Machado, combining traditional criticism and digital humanities. Using a multifaceted analysis encompassing lexical, semantic, and prosodic features, we quantify the stylistic divergences between the two brothers. Lexical analysis, including frequency and TF-IDF profiling, confirms Antonio's persistent focus on the Castilian landscape and temporal reflection, which contrasts with Manuel's aesthetic sensuality and cosmopolitan *Modernismo*. Psycholinguistic concreteness norms show that Antonio's diction is significantly more concrete, rooted in tangible, natural imagery, while Manuel's favors abstraction and ornamentation. Valence and transformer-based sentiment analyses further suggest Antonio's restrained emotional tone, in contrast to Manuel's broader affective range. Furthermore, automatic metrical analysis using Rantanplan identifies both poets' reliance on traditional Spanish meters (the octosyllable and hendecasyllable), yet highlights Antonio's gradual consolidation toward formal simplicity against Manuel's experimental variety. Grounded in statistical evidence, our findings frame Antonio's poetics of introspection and Manuel's poetics of aesthetic vitality as complementary responses to Spain's fin-de-siècle crisis. This research, therefore, demonstrates the validity of computational stylistics and proposes a robust, multi-layered model for digital literary analysis.

Keywords: Antonio Machado; Manuel Machado; lexical analysis; affective analysis; metrical analysis; *Modernismo*

Authors: Sunghyun Jang is Professor in the Department of English Language and Literature at Korea University. His areas of specialization are British Romantic poetry, Victorian poetry, and digital humanities (Email: sunghyun-j@korea.ac.kr). **Hye-Yoon Chung** is Associate Professor in the Department of Spanish Language and Literature at Korea University. Her areas of specialization are information structure, pragmatics in Spanish, and digital humanities (Email: hye_c@korea.ac.kr). This work was supported by the Ministry of Education of the Republic of Korea and the National Research Foundation of Korea (NRF-2023S1A5A8078836).

标题：词汇、情感与格律：安东尼奥·马查多与曼努埃尔·马查多诗学特质的定量研究

内容摘要：本研究采用数据驱动的研究框架，融合传统文学批评与数字人文方法，运用计算分析手段重新审视了安东尼奥·马查多与曼努埃尔·马查多兄弟的诗歌创作。从词汇、语义与格律三个维度出发，系统量化了两位诗人之间的风格差异。基于词频与 TF-IDF 文本特征的分析表明，安东尼奥始终聚焦于卡斯蒂利亚的景观与时间性沉思，这与其弟曼努埃尔追求感官美学和世界主义色彩的现代主义风格形成鲜明对照。基于心理学具体性规范的分析显示，安东尼奥所用词汇显著更具具象性，多依托于具体可感的自然意象；而曼努埃尔的语词则更倾向于抽象与修饰。情感效价与基于 Transformer 模型的情感分析进一步揭示，安东尼奥的情感基调偏向克制，而曼努埃尔的情感表达则更为丰富多元。此外，借助 Rantanplan 工具进行的自动格律分析显示，两位诗人都以西班牙传统格律（八音节诗与十一音节诗）为主，但安东尼奥明显趋向于形式上的简洁统一，而曼努埃尔则呈现出实验性的多样化风格。在实证统计数据的支撑下，本研究将安东尼奥内省的诗学与曼努埃尔美学活力的诗学，解读为对 19 世纪末西班牙危机的两种互补性回应。综上，本研究不仅验证了计算风格学在文学研究中的有效性，也为数字文学分析提供了一个稳健、多层次的模型框架。

关键词: 安东尼奥•马查多; 曼努埃尔•马查多; 词汇分析; 情感分析; 格律分析; 现代主义

作者简介：张圣贤，高丽大学英语语言文学系教授，主要研究领域为英国浪漫主义诗歌、维多利亚时代诗歌及数字人文；郑惠允，高丽大学西班牙语言文学系副教授，主要研究领域为信息结构、西班牙语语用学及数字人文。

1. Introduction

At the turn of the twentieth century, the Spanish literary world found itself at the heart of the turmoil of the *fin de siècle*. It was an era of grave crisis in national identity, as Spain was forced to reckon with its waning glory. The Disaster of ’98—Spain’s defeat in the Spanish-American War—instilled a widespread sense of helplessness and an urgent demand for introspection across the nation. It was during this critical transitional period that Manuel Machado and Antonio Machado emerged, both seeking to dismantle established conventions and capture a renewed national spirit. The literary divergence between the Machado brothers was, in essence, a manifestation of their different responses to the pivotal question of their era: Should art transcend reality, or should it probe more deeply into the heart of suffering?

The elder brother, Manuel Machado, was a leading figure of the Spanish *Modernismo* movement, heavily indebted to two major French currents.[1] The first was a preoccupation with formal perfection, inherited from Parnassianism. This approach treated a poem not as a spontaneous outpouring of emotion but as a meticulously crafted work of art. As a result, his poetry displays meticulous metrical control and structural stability, thus exuding a self-contained beauty. The second influence was the musicality and allusiveness of language, absorbed from Symbolism. Manuel demonstrated a remarkable gift for crafting the resonance of words, the allure of rhythm, and the interplay of the senses. These two influences coalesced into his literary creed of "art for art's sake," a declaration that sought to free art from any obligation to convey political messages or impart moral lessons.[2] This attitude was also shaped by the cosmopolitan sophistication he acquired while living in Paris.[3] Yet, paradoxically, while his poetry bore an international polish, its deepest roots were planted in Andalusia. The Andalusia he depicted was an aesthetically distilled festival of the senses. He infused his poetry with the spirit of *cante hondo*, i.e. the "deep song" of flamenco[4], and evoked exotic landscapes and fragments of bohemian life in dazzling colors and fluid rhythms. His poetic world is grounded in an aristocratic and refined sensibility that privileged the capture of moments of aesthetic beauty over the exploration of life's tragic depths.

While Manuel broadened the horizons of Spanish poetry through a world of vibrant sensory experience, his younger brother, Antonio Machado, is regarded as the most genuine voice of the Generation of '98—a writer who delved into the nation's deepest interior. The key framework for understanding Antonio's literary universe is the collective task that the Disaster of '98 imposed on its intellectuals. Dissatisfied with the aesthetic escapism of the prevailing *Modernismo*, Antonio and his fellow intellectuals of the Generation of '98 considered it their historical and

1 See Gordon Brotherston, *Manuel Machado: A Revaluation*, Cambridge: Cambridge University Press, 1968; Eloy Navarro Domínguez, "El 'Museo' de Manuel Machado," *Philologia Hispalensis* 1 (1994): 17-32; Rafael Alarcón Sierra, *Entre el modernismo y la modernidad: La poesía de Manuel Machado (Alma y Caprichos)*, Sevilla: Diputación Provincial de Sevilla, 1999.

2 See Monique Anne-Marie Cimon, *La poesía de Manuel Machado*, 1972, Florida State University, PhD dissertation; José A. Miranda Poza, "Rubén Darío, Manuel Machado, Antonio Machado: El modernismo hispánico a ambos lados del océano," *Literatura, História e Subjetividades: Perspectivas Latino-Americanas*, edited by Ana C. Alves et al., Teresina: EDUFPI, 2015, 113-135.

3 See Rafael Alarcón Sierra, "Contra Manuel Machado: La deconstrucción de su no-identidad en El mal poema," *Hispanic Review* 2 (2000): 161-181.

4 See Juan Vadillo, "La pena andaluza en Cante hondo de Manuel Machado," *Sinfonía Virtual* 41 (2021): 1-13.

ethical duty to confront Spain's wounds directly and seek a path toward renewal.[1] In service of this mission, Antonio rejected the ornamental splendor on display in his brother's work—the formal artifice and sensory indulgence.[2] He believed that true poetry was not a decorative play of language but a "word in time," a creation distilled from the poet's lived experience and gradually maturing thought.[3] His poetry is thus saturated with profound introspection and earnest contemplation of the condition of his country. The primary stage for Antonio's philosophical reflections was the Castilian plateau, particularly the landscape surrounding Soria. In this barren land, he perceived Spain's glorious past, its painful present, and the possibility of its regeneration. Through his engagement with the concrete landscapes of Castile, he expanded his meditations to encompass both the historical essence of Spain and the universal existential condition of humanity. Characterized by a restrained simplicity and depth, his poetry resonates powerfully with the reader.

To date, critics and academics have established the literary stature of both poets primarily through close readings, i.e., analyzing their philosophical foundations, thematic orientations, and aesthetic sensibilities. Such qualitative analyses have been indispensable for understanding the richness of their works. Nevertheless, within this extensive body of scholarship, there have been virtually no attempts to examine the linguistic features and stylistic differences of the brothers in a systematic, quantitative manner. Specifically, research that compares their entire corpora using methodologies drawn from the digital humanities remains unexplored.

The significance of the present study lies in its effort to fill this scholarly gap and to re-examine the poetic worlds of the Machado brothers through the lens of digital analysis. This study seeks to contribute to the field in two principal ways. First, it provides objective, data-driven evidence to complement and extend existing qualitative readings. This will enable discussions of the poets' stylistic differences to be grounded in concrete, empirical data. Second, it uncovers both macro-level patterns and micro-level linguistic habits that are difficult for the human eye to detect. By analyzing the complete poetic corpora of both brothers as a unified

1 See Elba Birmingham-Pokorny, "Campos de Castilla: An Insight into Antonio Machado's Poetic and Ideological Evolution," *Confluencia* 1 (1989): 17-24. For a broader account of the relationship between literature and historically situated ethical consciousness, see Nie Zhenzhao, "Ethical Literary Criticism: A Basic Theory," *Forum for World Literature Studies* 2 (2021): 189-207.

2 See M. Fernando Varela Iglesias, "Imaginación e intuición en la poética de Antonio Machado," *Revista Hispánica Moderna* 1 (2000): 59-74.

3 Machado articulated his mature poetics through his alter-ego, Juan de Mairena, famously defining poetry as "the essential word in time" (*la poesía es la palabra esencial en el tiempo*) (Machado, *Juan de Mairena*, 1936).

dataset, this study opens up new possibilities for identifying unexpected points of convergence and for highlighting critical divergences that have previously gone unnoticed. Ultimately, such an approach will not only refine our understanding of Antonio and Manuel but will also position this research as a pioneering case study in applying digital humanities methodologies to Spanish literary studies.

2. Methods

We compiled two author-specific corpora from authoritative critical editions. For Antonio Machado, the corpus includes all officially published poetry collections—*Soledades* (1899-1907), *Campos de Castilla* (1907-1917), *Nuevas canciones* (1917-1930), and *De un cancionero apócrifo* (1924-1936; poetic sections only)—as well as the *Poesías de la guerra* section (1936-1939), drawn from the RBA edition of *Obras completas I*[1]. In total, this corpus comprises 189 poems and 41,920 word tokens. For Manuel Machado, the corpus is based on his fourteen poetry books collected in *Poesías completas*[2]: *Alma* (1898-1900), *Caprichos (*1900-1905), *La fiesta nacional (Rojo y negro)* (1906), *El mal poema* (1909), *Museo* (1910)[3], *Apolo. Teatro pictórico* (1910), *Cante hondo* (1912), *Sevilla* (1920), *Poemas varios* (1921), *Ars moriendi* (1922), *Dedicatorias* (1910-1922), *Phoenix* (1935), *Horas de oro* (1938)[4], and *Cadencias de cadencias (Nuevas dedicatorias)* (1943). This corpus comprises 355 poems and 39,910 word tokens.

Following Shang and Underwood's distinction between semantic and prosodic features in poetry, this study compares the two poets' works along these dimensions.[5] On the semantic side, we first examined the most frequent lemmas across the main grammatical categories and identified words that distinguish the two poets. We then applied psycholinguistic norms, drawing on the LLM-derived Spanish word norms of Martínez et al.[6], which provide token-weighted averages of

1 See Antonio Machado, *Obras completas I*, Barcelona: RBA Editores, 2005.

2 See Manuel Machado, *Poesías completas,* Seville: Editorial Renacimiento, 2019.

3 In the current study, *Museo* (1910) refers to the independent collection included in the *Poesías completas* (M. Machado, 2019) edition, not to be confused with the section titled "Museo" within *Alma* (1898-1900) discussed in earlier criticism (e.g., Brotherston, 1968; Cimon, 1972; Navarro Domínguez, 1994).

4 *Horas de oro* (1938) consists largely of reused earlier poems compiled to project Machado's post-war patriotic and religious stance (Stauder, 1998). To avoid token duplication in the corpus, this study uses the *Poesías completas* (Renacimiento, 2019) edition, which retains only non-duplicated poems.

5 See Wenyi Shang and Ted Underwood, "Disentangling semantic and prosodic features of English poetry," *Digital Scholarship in the Humanities* 40.Supplement_1 (2025): i87-i99.

6 See Gonzalo Martínez et al., "AI-Generated Estimates of Familiarity, Concreteness, Valence, and Arousal for over 100,000 Spanish Words," *Quarterly Journal of Experimental Psychology* 1 (2024): 2272-2283.

concreteness (7-point scale: 1 = abstract, 7 = concrete) and valence (9-point scale: 1 = negative, 9 = positive). To assess statistical reliability, we employed Cohen's *d* (effect size), bootstrap 95% confidence intervals, and two-sided permutation tests. For the prosodic analysis, we employed Rantanplan[1], a rule-based pipeline for automatic scansion that performs POS tagging, syllabification, stress assignment, and metrical adjustment, while also identifying synalephas/synaereses and deriving metrical patterns. This enabled us to extract information on verse length, meter, and rhyme, which we then summarized by book. Following the case study on García Lorca by Benito-Santos et al.[2], we visualized the results using scansion-driven profiles, which include book-level meter distributions and diachronic heatmap profiles across collections.

This methodological design integrates semantic profiling and prosodic scansion within a unified framework. We combine frequency-based lexical analysis, psycholinguistic norms, and traditional stylistic indicators with automatic metrical detection and visualization in order to provide an integrated account of the Machado brothers' poetic styles. This dual approach ensures comparability across their poetic works and bridges digital stylistics with traditional literary scholarship. In doing so, it lays the groundwork for a more nuanced interpretation of their contrasting yet interconnected poetic projects.

3. Results

3.1 Lexical Analysis: Frequency and TF-IDF

3.1.1 Lemma Frequency and Lexical Category Distributions

The following analysis of the thirty most frequent lemmas (Fig. 1) in the poetic corpora of Antonio and Manuel Machado reveals a divergence in their primary thematic concerns and poetic worlds.

Although both poets share a foundational vocabulary, their most characteristic words point to two different sensibilities. Antonio's word frequency profile, for instance, confirms his established persona as a poet deeply connected to the Castilian landscape and the philosophical meditations it inspires. His most frequent lemmas such as *tierra* (earth, land), *mar* (sea), *campo* (field, countryside), *agua* (water), *sol* (sun), *viento* (wind), and *río* (river) anchor his poetry in a tangible geography. This

1 See Javier De la Rosa et al., "Rantanplan, Fast and Accurate Syllabification and Scansion of Spanish Poetry," *Procesamiento del Lenguaje Natural* 65 (2020) : 83-90; Guillermo Marco et al., "Automated Metric Analysis of Spanish Poetry: Two Complementary Approaches," *IEEE Access* 9 (2021): 51734-51746.

2 See Alejandro Benito-Santos et al., "Characterizing the Visualization Design Space of Distant and Close Reading of Poetic Rhythm," *Frontiers in Big Data* 6 (2023): article 1167708.

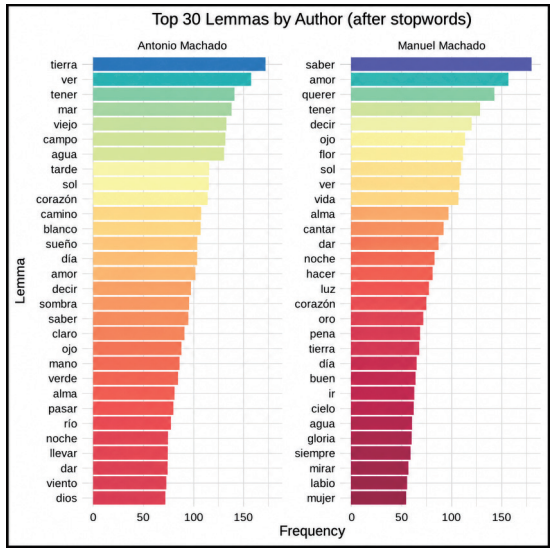


Fig. 1. Top 30 Lemmas in the Poetry of Antonio (left) and Manuel Machado (right)

lexical constellation embodies the Castilian landscape that Shaw[1] and Birmingham-Pokorny[2] described as central to the *Generación del 98*. Words like *viejo* (old), *tarde* (afternoon), *camino* (road, path), and *sombra* (shadow) are prominent, too. This vocabulary underscores his preoccupation with the passage of time, memory, and the existential journey of life. The motif of *camino* crystallizes Machado's meditation on temporality and human movement, echoing Lappin's[3] interpretation of the *Duero* landscape as a symbolic geography of contemplative introspection.

Manuel's vocabulary evokes a bohemian aesthetic of sensuous beauty, consistent with his association with *Modernismo* and the cosmopolitan artistic circles of his time. His list is topped by words of feeling and desire: *saber* (to know, but also to taste), *amor* (love), and *querer* (to want, to love). The presence of *mujer* (woman), *labio* (lip), and *pena* (sorrow) highlights his sustained attention to human intimacy and affective expression. In addition, aesthetic terms such as *flor* (flower), *cantar* (to sing), *oro* (gold), and *gloria* (glory) reveal a sustained preoccupation with art, beauty, and performance. His vocabulary is more decorative and outwardly expressive than his brother's. Overall, Manuel's lexicon suggests a poetry attentive to sensation and the fleeting, vibrant moments of life, love, and art. As observed

1 See Donald Shaw, *La generación del 98*, Madrid: Ediciones Cátedra, 1982.

2 See Elba Birmingham-Pokorny, "Campos de Castilla: An Insight into Antonio Machado's Poetic and Ideological Evolution," *Confluencia* 1 (1989): 17-24.

3 See Anthony John Lappin, "A orillas del Duero: Machado's Walk through the Country," *Studia Neophilologica* 1 (2012): 33-46.

in Rodríguez Martínez[1] and Alarcón Sierra[2], this lexicon reflects the *modernista* synthesis of sound, color, and rhythm that Miranda Poza[3] attributes to Rubén Darío's influence.

As illustrated in Fig. 2, although the brothers share a significant portion of their core vocabulary, the differences in their most-used verbs—as well as the subtle shifts in ranking—serve as quantitative evidence of their distinctive poetic sensibilities and thematic concerns.

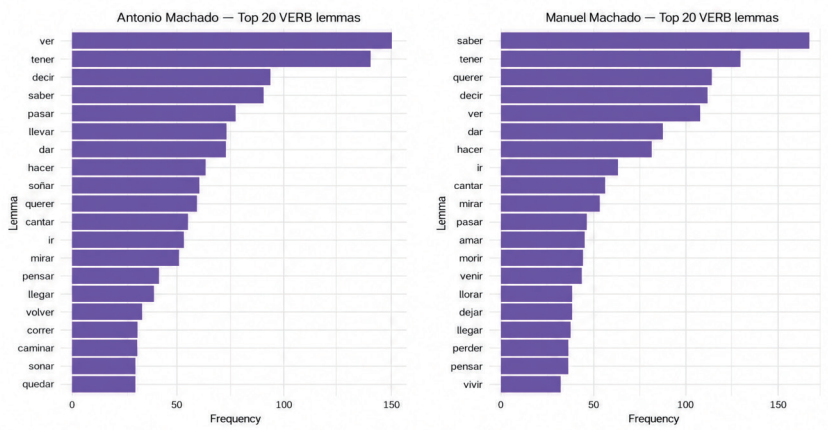


Fig. 2. Most Frequent Verbs in the Poetry of Antonio (left) and Manuel Machado (right)

A striking initial observation is the substantial overlap between the two lists. Thirteen verbs appear in the top twenty for both poets, including *ver* (to see), *saber* (to know), *tener* (to have), *decir* (to say), and *querer* (to want/love). Yet even within this common matrix, their ranking and thematic clusters diverge significantly. For Antonio, verbs of perception and motion—*ver*, *caminar* (to walk), *correr* (to run), *volver* (to return)—form the backbone of his poetic diction. This constellation of action and observation articulates what Lappin[4] describes as the poet-as-traveler motif, in which the journey becomes a spiritual act. The prominence of *ver* confirms

1 See Nora Rodríguez Martínez, "Modernismo en la poesía de Manuel Machado: Tradición e innovación en el soneto alejandrino," *Rhythmica: Revista Española de Métrica Comparada* 13 (2015): 193-209.

2 See Rafael Alarcón Sierra, "Contra Manuel Machado: La deconstrucción de su no-identidad en El mal poema," *Hispanic Review* 2 (2000): 161-181.

3 See José A. Miranda Poza, "Rubén Darío, Manuel Machado, Antonio Machado: El modernismo hispánico a ambos lados del océano," *Literatura, História e Subjetividades: Perspectivas Latino-Americanas*, edited by Ana C. Alves et al., Teresina: EDUFPI, 2015, 113-135.

4 See Anthony John Lappin, "A orillas del Duero: Machado's Walk through the Country," *Studia Neophilologica* 1 (2012): 33-46.

Antonio's orientation toward an outward gaze that becomes inward reflection, a pattern also identified in de Ros[1]. In Manuel's case, by contrast, the hierarchy reverses: *saber* (to know) and *amar* (to love) surpass *ver*, while *llorar* (to cry) also ranks prominently.

3.1.2 TF-IDF Approach

Figure 3 presents the analysis of TF-IDF signatures in the poetry of Antonio and Manuel. Far from being random lexical assortments, the two lists of high-scoring terms function as concise summaries of their divergent poetic projects.

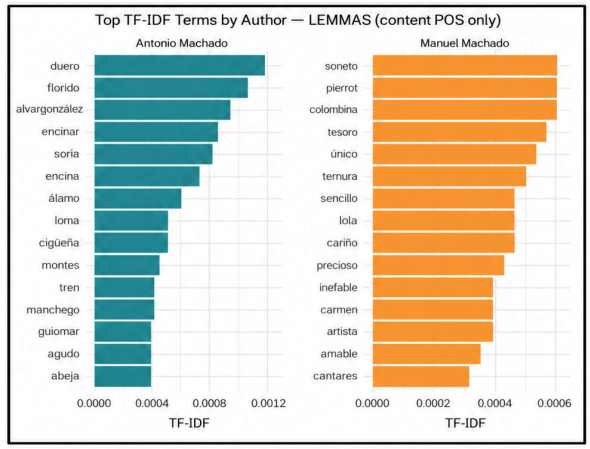


Fig. 3. Lexical Distinctiveness in the Poetry of Antonio and Manuel Machado (TF-IDF)

In Antonio's corpus, the distinctive vocabulary is embedded in the geography of Castile, the symbolic heartland of Spain. His top terms—*Duero*, *florido* (flowery), *Alvargonzález*, *encinar* (oak grove), *Soria*, *encina* (holm oak), *álamo* (poplar), *loma* (hill), *montes* (mountains), and *abeja* (bee)—form a lexicon of landscape and reflection. *Duero*, the river that runs through *Soria*, is not merely a place name but a recurrent emblem of time, memory, and consciousness. Supporting items such as *encinar*, *encina*, and *álamo* anchor his poetry in an elemental natural world that sustains philosophical meditation on being and change.[2] Even the apparently decorative *florido* retains this resonance. The appearance of *Alvargonzález* points directly to *La tierra de Alvargonzález*, the long narrative poem where human experience unfolds as an allegory of fate and memory within the Soria landscape.[3]

1 See Xon De Ros, *The Poetry of Antonio Machado: Changing the Landscape*, Oxford: Oxford University Press, 2015.

2 See Enrique Gallud Jardiel, "Antonio Machado: Conocimiento y lógica," *Endoxa: Series Filosóficas* 42 (2018): 389-408.

3 See Ricardo Gullón, *Una poética para Antonio Machado*, Madrid: Espasa Calpe, 1970.

Within the broader framework of the Generation of '98, Antonio's Castile exemplifies the ideal of *paisaje como conciencia*—landscape as consciousness—identified by Shaw.[1] Antonio's TF-IDF profile thus confirms that geography itself becomes an instrument of thought.

In contrast, Manuel's list of distinctive terms evokes a world of aesthetic awareness and sensory performance. His top words construct a lexicon centered on art, emotion, and performance. Specifically, *soneto* (sonnet) being a top term indicates a focus on poetic form. *Pierrot* and *Colombina*, characters from the c*ommedia dell'arte*, signal a departure from Spanish soil into a cosmopolitan, decadent European aesthetic. The word *artista* (artist) further confirms a self-conscious preoccupation with the creator's role. This diction constructs a *modernista* verbal tableau fusing sound and color[2], echoing Juan Ramón Jiménez's[3] account of *modernista* musicality. Whereas Antonio has the real Castile, Manuel presents a sensual and aestheticized vision of Andalusia. This is evident in terms like *Carmen* and *Lola*—archetypal female names associated with flamenco and passion—and *cantares* (folkloric songs). It is an Andalusia seen through the lens of art and music, rather than as a lived geographical reality. Emotional and evaluative terms such as *tesoro* (treasure), *único* (unique), *ternura* (tenderness), *cariño* (affection), *precioso* (precious), and *inefable* (ineffable) reinforce this expressive tendency, highlighting a lyricism attuned to sensory nuance and verbal texture. His TF-IDF lexicon, therefore, encapsulates a poetics of form and emotion, complementing Antonio's focus on place and perception. Yet, in Manuel's later collections— *Phoenix* (1935), *Horas de oro* (1938) and *Cadencias de cadencias* (1943)—his voice acquires a markedly different inflection: devotional, patriotic, and elegiac. The lexicon of these works incorporates terms of faith and transcendence (*alma*, *cielo*, *Dios*), alongside civic and national imagery that resonates with Spain's postwar rhetoric.

3.2 Semantic and Affective Dimensions

Building on the lexical findings discussed above, this section turns to the semantic and affective dimensions of the Machado brothers' poetry. Beyond the frequency of specific lemmas, these psycholinguistic variables capture how each poet's language shapes a world of perception and emotion. *Concreteness* and *valence* provide complementary perspectives on this issue. Concreteness measures

1 See Donald Shaw, *La generación del 98*, Madrid: Ediciones Cátedra, 1982.

2 See Eloy Navarro Domínguez, "El 'Museo' de Manuel Machado," *Philologia Hispalensis* 1 (1994): 17-32; Nora Rodríguez Martínez, "Modernismo en la poesía de Manuel Machado: Tradición e innovación en el soneto alejandrino," *Rhythmica: Revista Española de Métrica Comparada* 13 (2015): 193-209.

3 See Juan Ramón Jiménez and Jorge Urrutia, *El modernismo: Apuntes de curso (1953)*, Madrid: Visor, 1999.

the extent to which poetic language refers to concepts that can be experienced through interaction with the physical or sensory world. Valence, in turn, reflects the affective tone evoked by words, ranging from negative to positive affect.

The following subsections examine, first, the distribution of concreteness and then that of valence, to illuminate how perception and affect coalesce in the linguistic texture of Antonio's and Manuel Machado's verse.

3.2.1 Analysis of Concreteness

Following Martínez et al.[1], concreteness measures the extent to which a word denotes a concept that can be experienced through interaction with the physical environment; it thus quantifies how directly poetic language evokes sensory experience. Given the traditional contrast between the two brothers—Antonio's reflective austerity and Manuel's decorative *Modernismo*—one might expect Manuel's lexicon, rich in color and emotion, to score higher in concreteness, and Antonio's more meditative diction to appear comparatively abstract. The following analysis tests this assumption by comparing their complete poetic corpora. It asks whether Antonio's philosophical verse is, in fact, less sensory—or whether his engagement with landscape grounds it more firmly in the tangible world.

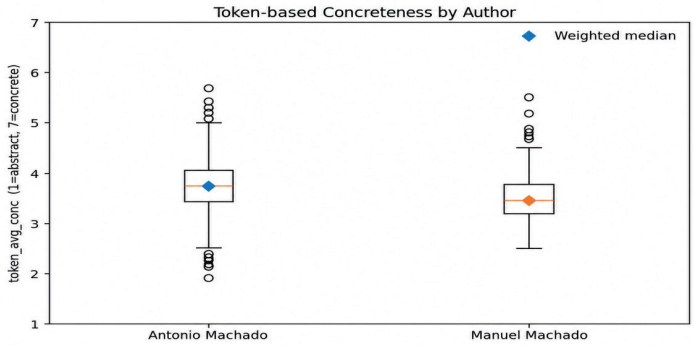


Fig. 4. Concreteness in the Poetry of Antonio and Manuel Machado
(Token-weighted; 7-point Scale: 1 = Abstract, 7 = Concrete)

Figure 4 presents a token-weighted analysis, which measures the average concreteness of all word tokens in the text. This approach gives more weight to words that are used frequently. The analysis reveals that Antonio's work consistently scores higher on the concreteness scale. His token-weighted average concreteness score is approximately 3.757, compared to Manuel's 3.508. This results in a significant difference of about 0.249 points.

1 See Gonzalo Martínez et al. "AI-Generated Estimates of Familiarity, Concreteness, Valence, and Arousal for over 100,000 Spanish Words," *Quarterly Journal of Experimental Psychology* 1 (2024): 2272-2283.

Table 1 Statistical Summary of Concreteness Measures (Token-weighted)

	Antonio Machado (*M*)	Manuel Machado (*M*)	Difference (Δ*M*)	Cohen's	95% CI [LL, UL]	-value
Token-weighted	3.7569	3.5075	0.2493	0.6617	[0.1599, 0.3614]	0.0207

First, Cohen's *d* yielded a value of 0.66, which is considered a moderate effect. This suggests that the difference in concreteness between the poets is substantial. Second, the resulting bootstrap 95% confidence interval (CI) [0.160, 0.361] is entirely above zero. This provides strong evidence that Antonio's poetry is more concrete than Manuel's and that this result is not a product of random chance. Finally, a permutation test, which determines the probability of observing such a difference by chance, resulted in a *p*-value of 0.0207, well below the standard alpha level of 0.05. These tests confirm that the observed difference in concreteness is statistically significant.

The divergence in concreteness between the Machado brothers attests to their differing artistic projects. Antonio's work is famously rooted in the landscapes of Castile, as exemplified by his collection *Campos de Castilla*. He grounds his existential and temporal meditations in tangible, everyday elements. Although Antonio's poetry is often described as reflective and philosophical, his diction remains grounded in the tangible world of nature and landscape. This reliance on sensory vocabulary explains his higher concreteness scores, as perceptible elements of the physical environment become a philosophical medium.

Manuel's lower concreteness score (and thus higher abstraction) is characteristic of his affiliation with *Modernismo*. This movement prized aestheticism, musicality, and a more ornate, cosmopolitan vocabulary. Consequently, his language is more likely to feature abstract concepts and words chosen for their sound and elegance rather than their direct connection to the physical environment. Although his verse abounds in sensory imagery, these words function less as depictions of the physical world than as aesthetic symbols of emotion, color, and sound. In line with *Modernismo*, his diction transforms sensory elements into vehicles of expression, producing a more abstract and self-consciously artistic lexicon.

3.2.2 Dual-Method Analysis of Affect

The affective dimension of the Machado brothers' poetry, long noted in literary criticism, has seldom been examined through quantitative means. Antonio is often portrayed as a poet of restraint and quiet melancholy, grounding emotion in the landscapes of Castile. Manuel, in contrast, embodies the expressive and ornamental sensibility of *Modernismo*, marked by rhythm, sensuality, and aesthetic

self-awareness. While such contrasts between *seriedad* and *ornamento* have been described qualitatively, this study seeks to assess them empirically by analyzing the emotional tone of their poetic language.

3.2.2.1 Lexicon-Based Valence

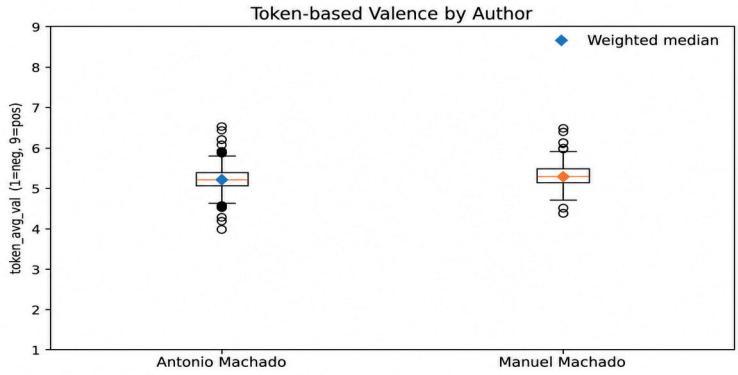


Fig. 5. Valence in the Poetry of Antonio and Manuel Machado
(Token-weighted; 9-Point Scale: 1 = Negative, 9 = Positive)

As illustrated in Fig. 5, the analysis reveals a subtle but consistent difference in the emotional valence of the two poets' works. Overall, Manuel's poetry exhibits a slightly more positive emotional tone than Antonio's. The token-weighted mean valence for Antonio's corpus is 5.2336, whereas Manuel's is marginally higher at 5.3146. The box plot visualizes this distinction: Manuel's entire distribution is shifted slightly upward on the 9-point scale compared to Antonio's. Although subtle, this difference supports the critical view that Manuel's poetic voice, influenced by *Modernismo*, favors a more decorative and rhythmically expressive style, whereas Antonio's diction is marked by concision and essential clarity.

Table 2. Statistical Summary of Valence Measures (Token-weighted)

	Antonio Machado (*M*)	Manuel Machado (*M*)	Difference (Δ*M*)	Cohen's	95% CI [LL, UL]	-value
Token-weighted	5.2336	5.3146	-0.0810	-0.3942	[-0.1272, -0.0158]	0.1355

Statistically, the direction of difference is stable (Cohen's $d \approx -0.39$; 95% CI [–0.127, –0.016]), although the permutation test ($p \approx 0.136$) indicates that the difference does not reach statistical significance. Even so, the trend supports earlier qualitative insights: Antonio's poetry converts existential meditation into emotionally neutral or subdued expression, whereas Manuel's lyricism leans toward exuberance and aesthetic display.

3.2.2.2 Transformer-Based Sentiment

To contextualize these lexical findings, a transformer-based sentiment model (*Pysentimiento RoBERTa*) was applied to both corpora. Unlike word-based valence estimation, this approach classifies each poetic unit as positive, negative, or neutral based on contextual semantics and composition. At the corpus level (Fig. 6), both poets show broadly balanced sentiment distributions, with neutral tone exceeding 45% of the data. Antonio's poems contain a higher proportion of negative sentiment (≈ 33%) than Manuel's (≈ 28%); this confirms the image of Antonio as a poet of restrained melancholy whose work oscillates between compassion and resignation. Manuel's slightly greater share of positive sentiment (≈ 26%) reflects his *modernista* delight in verbal artifice and expressive form.



Fig. 6. Sentiment Distribution across the Poetry of Antonio and Manuel Machado (via *Pysentimiento*)

Collection-level results (Fig. 7 and Fig. 8) reveal clear intra-author contrasts. In Antonio's oeuvre, *Campos de Castilla* and *Poesías de la Guerra* register the highest negative proportions (≈ 33–37%), and *Nuevas canciones* and *De un cancionero apócrifo* present more balanced ratios, marking his late shift toward introspection.

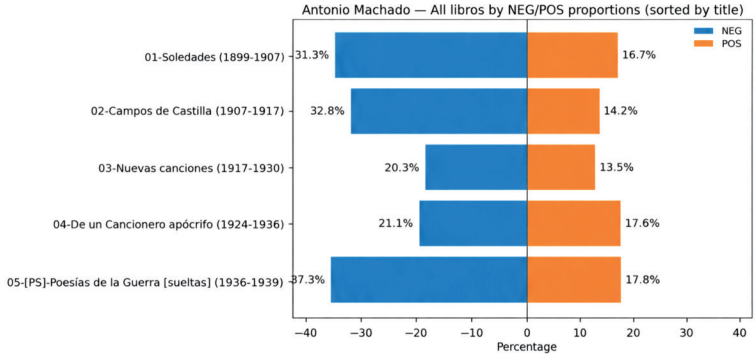


Fig. 7. Sentiment Distribution across Antonio Machado's Collections (via *Pysentimiento*)

In Manuel's corpus, *Cadencias de cadencias* (1943) and *Horas de oro* (1938) exhibit the highest positivity (≈ 33–35%), whereas *La fiesta nacional* and *El mal poema* display more negativity (≈ 36%). These fluctuations reflect Manuel's stylistic range—from ironic detachment to lyrical exuberance—within a single aesthetic framework.

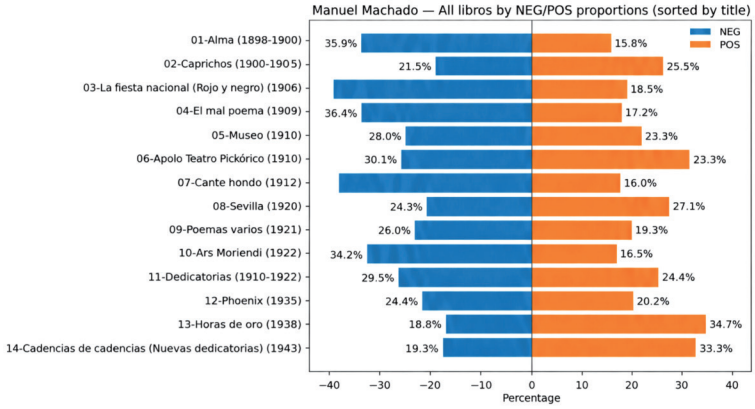


Fig. 8. Sentiment Distribution across Manuel Machado's Collections (via *Pysentimiento*)

When interpreted together, the lexicon-based and transformer-based results indicate complementary tendencies. Both approaches point to Antonio's consistent emotional restraint and to Manuel's broader expressive range, findings that align with—without establishing—the contrasts often noted in literary criticism. The markedly higher proportion of neutral classifications in Antonio's corpus provides a plausible explanation for the divergence observed between lexical and contextual sentiment measures. In the lexicon-based analysis, the predominance of emotionally neutral or descriptive words—terms associated with landscape, temporality, and introspection—drives his mean valence toward the midpoint of the scale (≈ 5.2). By contrast, the transformer-based sentiment model, which interprets contextual meaning rather than individual word polarity, often reads these same passages as affectively negative, given their tone of stillness, solitude, or resignation. In other words, Antonio's language tends to encode affect implicitly rather than lexically, whereas Manuel's poetry displays a lower share of neutral instances and a closer alignment between the two analytical approaches, reflecting a more explicit emotional register and a wider expressive range characteristic of *Modernismo*.

3.3 Digital Prosody: A Study of Metrical Patterns

This section examines the metrical composition of the poetic corpora of the Machado brothers. The analyses that follow present findings from three

visualizations: an overall comparison of meter frequencies and two diachronic heatmaps showing the metrical evolution of each poet's work, volume by volume.

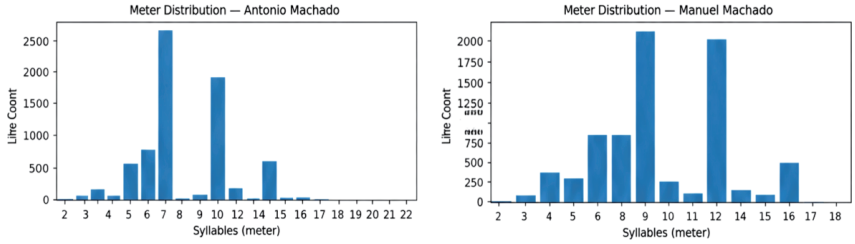


Fig. 9. Overall Meter Distribution in the Poetry of Antonio (left) and Manuel Machado (right)

Figure 9 shows that both Antonio and Manuel build their poetic practices around two foundational meters of the Hispanic tradition: the octosyllable (8 syllables) and the hendecasyllable (11 syllables). For Antonio, there is a clear reliance on these two forms. The octosyllable—the quintessential meter of traditional Spanish ballads (*romances*)—is his most frequently employed verse form, appearing over 2,500 times. This is followed closely by the hendecasyllable, the classic meter of Renaissance art poetry, later revitalized by *Modernismo*, which he uses in approximately 2,000 lines. All other meters occur at much lower frequencies. Similarly, Manuel's poetry centers on the octosyllable and hendecasyllable, with each appearing around 2,000 times. However, a key distinction emerges in his broader metrical palette: hexasyllables (6) and heptasyllables (7) are markedly more frequent, and the alexandrine (14)—prominent within French Symbolism and Latin-American *Modernismo*—is a visible component of his corpus, while it remains minor in Antonio's. This overview suggests that while both poets are rooted in the same poetic tradition, Manuel exhibits a greater tendency toward metrical variety and experimentation.[1] These results align with traditional metrical studies, where Antonio's formal restraint[2] contrasts with Manuel's *modernista* rhythmic flexibility[3].

Moving beyond these qualitative readings, our corpus-wide analysis quantifies

1 See Manuel Romero Luque, "El soneto modernista (Manuel Machado como paradigma)," *Rhythmica: Revista Española de Métrica Comparada* 15 (2017): 113-145.

2 See Tomás Navarro Tomás, "La versificación de Antonio Machado," *Los poetas en sus versos: Desde Jorge Manrique a García Lorca*, edited by Tomás Navarro Tomás. Barcelona: Ariel, 1973, 273-289.

3 See Solita Salinas de Marichal, "Antonio y Manuel Machado: Tradición y originalidad," *Boletín de la Institución Libre de Enseñanza* 38 (1998): 19-37; José A. Miranda Poza, "Rubén Darío, Manuel Machado, Antonio Machado: El modernismo hispánico a ambos lados del océano," *Literatura, História e Subjetividades: Perspectivas Latino-Americanas*, edited by Ana C. Alves et al., Teresina: EDUFPI, 2015, 113-135.

their entire metrical output. By systematically measuring metrical frequencies across all officially published collections, the present study confirms that Antonio's verse remains consistently rooted in a narrow classical framework, and that Manuel's displays sustained formal variability and experimentation. In doing so, it translates long-standing interpretive generalizations—once grounded in intuition and selective textual evidence—into statistically verifiable patterns that reveal the structural foundations of the Machado brothers' rhythmic identities.

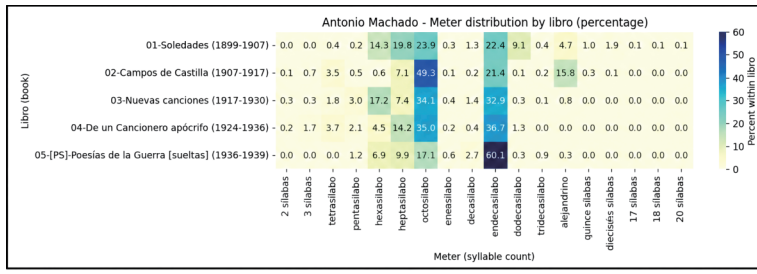


Fig. 10. Meter Distribution across Antonio Machado's Collections (Percentage by Book)

Figure 10 presents the diachronic heatmap of Antonio's poetic corpus, quantifying with precision the stylistic evolution long noted by critics. *Soledades* (1899-1907) displays marked metrical diversity—heptasyllables (19.8%), octosyllables (23.9%), and hendecasyllables (22.4%)—which mirrors the flexible musical rhythm and introspective tone described by Vila-Belda (2009). A clear reorientation appears in *Campos de Castilla* (1907-1917), where the octosyllable becomes dominant (49.3%), the hendecasyllable recedes (21.4%), and the alexandrine rises (15.8%). This shift confirms that Machado's verse turns from symbolist inwardness toward a plainer and more descriptive mode. The prevalence of shorter meters, rather than implying simplicity, suggests a conscious incorporation of popular cadence into his reflective style. In *Nuevas canciones* (1917-1930), Antonio restores balance between octosyllables (34.1%) and hendecasyllables (32.9%); he combines lyrical movement with aphoristic concision—a development consistent with his pursuit of structural coherence. A similar equilibrium persists in *De un cancionero apócrifo* (1924-1936), where both meters remain nearly equal in frequency (36.7% versus 35.0%); this statistical balance suggests a deliberate search for compositional unity within metrical restraint during his later years. Finally, *Poesías de la Guerra* (1936-1939) shows a pronounced concentration on the hendecasyllable (60.1%), which confirms Whiston's[1] view of the *Civil War*

1 See James Whiston, "'Más fuerte que la guerra': The Civil-War Sonnets of Antonio Machado," *The Modern Language Review* 3 (1993): 644-665.

Sonnets as the culmination of Antonio's tension between emotional intensity and formal control. In this final phase, meter serves as a principle of balance—a means of sustaining order and clarity amid historical disruption.

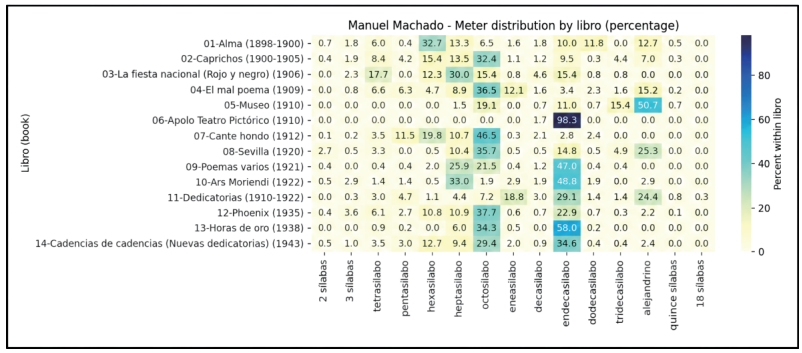


Fig. 11. Meter Distribution across Manuel Machado's Collections (Percentage by Book)

Unlike Antonio's, Manuel's metrical trajectory, as shown in Fig. 11, is characterized by discontinuity and experimentation. Each collection functions as a self-contained rhythmic project rather than a stage in a linear evolution. The diachronic data show that *Alma* (1898-1900) privileges short meters—particularly the hexasyllable—combined with longer verse forms (12-14 syllables), creating a texture of alternating rhythmic tension and release. This balance between brevity and musical expansion captures the *modernista* fusion of rhythm and melody, echoing the musicality of Rubén Darío. A turning point appears in *El mal poema* (1909), which, contrary to earlier assumptions about his preference for the hendecasyllable, relies mainly on octosyllables and mid-length meters. This pattern aligns with Alarcón Sierra's[1] observation of the poet's use of colloquial diction and self-referential tone as structural correlates of rhythmic irregularity. In *Apolo. Teatro pictórico* (1910), the poet turns almost exclusively to the hendecasyllable (98.3%); he transforms metrical uniformity into a principle of theatrical decorum and compositional harmony. This disciplined design reflects *Parnassian* concerns with symmetry and visual structure. Conversely, in *Cante hondo* (1912), the octosyllable becomes the rhythmic core, echoing Andalusian song forms and vernacular lexicon, as discussed in Vadillo[2]. His mid-career collection, *Ars moriendi* (1922), initiates a return to the hendecasyllable, a trend that becomes definitively established in his

1 See Rafael Alarcón Sierra, *Entre el modernismo y la modernidad: La poesía de Manuel Machado (Alma y Caprichos)*, Sevilla: Diputación Provincial de Sevilla, 1999.

2 See Juan Vadillo, "La pena andaluza en Cante hondo de Manuel Machado," *Sinfonía Virtual* 41 (2021): 1-13.

late collections—*Horas de oro* (1938) and the postwar *Cadencias de cadencias* (1943). This renewed concentration points to a return to measured form and tonal restraint.

4. Conclusion

This study set out to re-examine the poetic worlds of Antonio and Manuel Machado through a systematic, quantitative lens; it aimed to bridge the gap between traditional literary criticism and data-driven digital humanities approaches. By integrating lexical, semantic, affective, and prosodic analyses, we have provided empirical evidence that both confirms and refines the long-standing view of the brothers as two foundational yet divergent figures in modern Spanish poetry. Our findings reveal that their differing artistic visions are embedded not only in their thematic choices but in the very statistical texture of their language and verse.

Lexically and semantically, Antonio's poetry is anchored in concreteness and in the sensory immediacy of the Castilian landscape. His consistently higher concreteness scores and the recurrence of landscape-related nouns reveal a poetics that transforms material perception into philosophical meditation. Manuel, by contrast, privileges abstraction and ornamental expressivity; his lexicon is populated with words of emotion, art, and aesthetic refinement. Although the affective valence of their vocabularies differed only slightly, transformer-based sentiment analysis uncovered an important distinction: Antonio tends to encode emotion implicitly through neutral description, whereas Manuel favors a more explicit and expressive emotional register. These findings transform qualitative intuitions into quantifiable stylistic features.

A similar pattern emerges in the prosodic domain. While both poets built their work upon the foundational octosyllabic and hendecasyllabic meters of the Hispanic tradition, their approaches to form were markedly different. The diachronic analysis revealed Antonio's metrical practice to be an evolutionary trajectory. He moved from the youthful diversity of *Soledades* to a deliberate mastery and eventual consolidation of the hendecasyllable, forging a stable, mature style that mirrored his deepening philosophical inquiry. Manuel's career, conversely, was one of eclectic, collection-specific experimentation. His was the path of a protean stylist; his brother's that of a steadfast sage.

Yet we wish to emphasize that the Machado brothers emerge not as simple opposites but as parallel architects of Spanish modern lyricism: Antonio forging a poetics of introspection and temporal consciousness, and Manuel crafting one of aesthetic vitality and performance. Their differences—now mapped through

linguistic and rhythmic data—trace two different yet interwoven responses to the cultural crisis of early twentieth-century Spain.

Two limitations of this study suggest immediate directions for future research. First, the lexical norms and sentiment models employed, though broad in coverage, are not trained on the specialized registers of poetic language. Adapting or fine-tuning these models on annotated verse could significantly enhance their sensitivity to figurative language. Second, our unit of analysis remained at the poem and book levels. Incorporating phonetic and phonological features (e.g., assonance, alliteration) would offer a deeper understanding of how sound and sense interact in each poet's work. Beyond these limitations, the analytical framework established here opens several avenues for broader inquiry. A logical next step would be to extend this stylistic investigation to the brothers' theatrical collaborations. Applying author-attribution models to their co-written plays could reveal whether the different poetic signatures identified here persist in dramatic dialogue. Furthermore, this study could be expanded to situate the brothers within a broader network of their contemporaries. A comparative analysis involving key figures such as Rubén Darío, Juan Ramón Jiménez, and Miguel de Unamuno could quantitatively map these authors within a shared stylistic space. This may lead us to discover new relationships that challenge traditional literary groupings.

We believe that this study contributes meaningfully to scholarship on the Machado brothers by translating long-standing critical intuitions into verifiable, quantitative patterns. By demonstrating the efficacy of an integrated computational framework, this research also stands as a case study in the productive application of digital-humanities methodologies to Spanish literary analysis. Viewed through the empirical lens of data, the poetic legacies of Antonio and Manuel Machado emerge as two complementary and equally vital forces shaping the poetic language of Spanish modernity.

Works Cited

Alarcón Sierra, Rafael. "Contra Manuel Machado: La deconstrucción de su no-identidad en *El mal poema*." *Hispanic Review* 2 (2000): 161-181.

—. *Entre el modernismo y la modernidad: La poesía de Manuel Machado (Alma y Caprichos).* Sevilla: Diputación Provincial de Sevilla, 1999.

Benito-Santos, Alejandro et al. "Characterizing the Visualization Design Space of Distant and Close Reading of Poetic Rhythm." *Frontiers in Big Data* 6 (2023): article 1167708.

Birmingham-Pokorny, Elba. "*Campos de Castilla*: An Insight into Antonio Machado's Poetic and Ideological Evolution." *Confluencia* 1 (1989): 17-24.

Brotherston, Gordon. *Manuel Machado: A Revaluation.* Cambridge: Cambridge UP, 1968.

Cimon, Monique Anne-Marie. *La poesía de Manuel Machado.* 1972. Florida State University, PhD dissertation.

De la Rosa, Javier et al. "Rantanplan, Fast and Accurate Syllabification and Scansion of Spanish Poetry." *Procesamiento del Lenguaje Natural* 65 (2020): 83-90.

De Ros, Xon. *The Poetry of Antonio Machado: Changing the Landscape.* Oxford: Oxford UP, 2015.

Gallud Jardiel, Enrique. "Antonio Machado: Conocimiento y lógica." *Endoxa: Series Filosóficas* 42 (2018): 389-408.

Gullón, Ricardo. *Una poética para Antonio Machado.* Madrid: Espasa Calpe, 1970.

Jiménez, Juan Ramón and Jorge Urrutia. *El modernismo: Apuntes de curso (1953).* Madrid: Visor, 1999.

Lappin, Anthony John. "A orillas del Duero: Machado's Walk through the Country." *Studia Neophilologica* 1 (2012): 33-46.

Machado, Antonio. *Obras completas I*. Barcelona: RBA Editores, 2005.

Machado, Manuel. *Poesías completas.* Seville: Editorial Renacimiento, 2019.

Marco, Guillermo et al. "Automated Metric Analysis of Spanish Poetry: Two Complementary Approaches." *IEEE Access* 9 (2021): 51734-51746.

Martínez, Gonzalo et al. "AI-Generated Estimates of Familiarity, Concreteness, Valence, and Arousal for over 100,000 Spanish Words." *Quarterly Journal of Experimental Psychology* 1 (2024): 2272-2283.

Miranda Poza, José A. "Rubén Darío, Manuel Machado, Antonio Machado: El modernismo hispánico a ambos lados del océano." *Literatura, História e Subjetividades: Perspectivas Latino-Americanas*, edited by Ana C. Alves et al. Teresina: EDUFPI, 2015. 113-135.

Navarro Domínguez, Eloy. "El 'Museo' de Manuel Machado." *Philologia Hispalensis* 1 (1994): 17-32.

Navarro Tomás, Tomás. "La versificación de Antonio Machado." *Los poetas en sus versos: Desde Jorge Manrique a García Lorca*, edited by Tomás Navarro Tomás. Barcelona: Ariel, 1973. 273-289.

Nie Zhenzhao. "Ethical Literary Criticism: A Basic Theory." *Forum for World Literature Studies* 2 (2021): 189-207.

Rodríguez Martínez, Nora. "Modernismo en la poesía de Manuel Machado: Tradición e innovación en el soneto alejandrino." *Rhythmica: Revista Española de Métrica Comparada* 13 (2015): 193-209.

Romero Luque, Manuel. "El soneto modernista (Manuel Machado como paradigma)." *Rhythmica: Revista Española de Métrica Comparada* 15 (2017): 113-145.

Salinas de Marichal, Solita. "Antonio y Manuel Machado: Tradición y originalidad." *Boletín de la Institución Libre de Enseñanza* 38 (1998): 19-37.

Shang, Wenyi and Ted Underwood. "Disentangling semantic and prosodic features of English poetry." *Digital Scholarship in the Humanities* 40, Supplement 1 (2025): i87-i99.

Shaw, Donald. *La generación del 98.* Madrid: Ediciones Cátedra, 1982.

Stauder, Thomas. "La repercusión del 18 de julio de 1936 en la vida y la obra de Manuel Machado."

Vencer no es convencer: Literatura e ideología del fascismo español, edited by Mechthild Albert. Madrid: Iberoamericana Vervuert, 1998. 199-220.

Vadillo, Juan. "La pena andaluza en *Cante hondo* de Manuel Machado." *Sinfonía Virtual* 41 (2021): 1-13.

Varela Iglesias, M. Fernando. "Imaginación e intuición en la poética de Antonio Machado." *Revista Hispánica Moderna* 1 (2000): 59-74.

Vila-Belda, Reyes. "Antonio Machado y la creación de su prestigio literario." *Revista de Estudios Hispánicos* 2 (2009): 227-250.

Whiston, James. "'Más fuerte que la guerra': The Civil-War Sonnets of Antonio Machado." *The Modern Language Review* 3 (1993): 644-665.