

Pia Bhatia

Introduction to Digital Humanities

Professor Wouter Haverals

Spring 2024

Metaphor as Protest: A Close and Distant Look at Novels in Francoist Spain

Francisco Franco ruled Spain during the years 1936-1975 after coming to power post the Spanish Civil War via his role as a military leader of the Spanish Arm. Franco's decades-long regime, often referred to as *Franquismo*, was characterized by oppressive fascist policies and strict censorship.¹ Particularly noteworthy is the use of censorship during this period: no books could be published prior to being approved by the Ministry of Information. The list of unacceptable subject matter, even in works of fiction, was exhaustive, ranging from criticism of the government, military and police, to the "documentation of poverty, ignorance, economic stagnation, social injustice" as well as "writings favoring divorce, abortion, or euthanasia, justifying adultery, premarital eroticism, sexual license, prostitution..."² The results of this censorship—its distortion of history and cultural harms—are felt even today.

That said, many historians view this period retrospectively as fertile literary ground for novels that transgressed censorship in notably oblique ways. As Margaret Gonzalez writes: "Due to its subtlety, the political novel may provide the ideal expression of the politics of a culture or society. It is able to provide access to an era, offering a unique perspective of a specified time

¹ Perez, "Fascist Models and Literary Subversion," 73–74.

² Perez, "Fascist Models and Literary Subversion," 73–74.

and place. This aspect of the novel is of particular significance in studying the political environment of Spain, for as George Wellwarth notes, “The educated Spanish reader, . . . has been well described as a man who gets all his information by reading between the lines.””³ There is now extensive literature that analyzes the novels that were published during this period and looks at the ways in which they use figurative language to make bold political statements. That said, the works that scholars have chosen are not exhaustive; understandably, they pertain to the most popular novels during the period, as well as those with the greatest “staying power” over time.

I would like to propose a digital humanities project that engages with this body of work on multiple fronts. The most important aspect of the project will follow from Chris Tanasescu, Vaibhav Kesarwani, Diana Inkpen’s work on metaphor detection by deep learning in the digital humanities.⁴ The project will engage with five major works published under Franco: *La colmena* (The Hive, 1953) and *La familia de Pascual Duarte* (The Family of Pascual Duarte 1942), written by Nobel Prize winner Camilo Jose Cela, *Primera memoria* (School of the Sun, 1959) by Ana Maria Matute, *Cinco horas con Mario* (Five Hours with Mario, 1966) by Miguel Delibes, and *Escuadra hacia la muerte* (The Condemned Squad, 1953) by Alfonso Sastre.⁵ My reason for choosing these works is that they have already been analyzed prior by scholar Margaret Carmell Gonzalez, who has justified their relevance and representativeness as a range of resistance novels published during this time. Using the aforementioned researchers’ methodologies, the project will assess the uses of coded language (in the form of allegories, symbolism etc.). It both aims to

³ Gonzalez, “Literature of Protest,” 3.

⁴ Tanasescu, Kesarwani, and Inkpen, “Metaphor Detection by Deep Learning and the Place of Poetic Metaphor in Digital Humanities,” 1.

⁵ Gonzalez, “Literature of Protest,” 13.

contribute to literary studies by offering new perspectives on the works of authors from Francoist Spain, enriching interpretations of their thematic concerns and narrative strategies and to develop and refine digital humanities methodologies for metaphor analysis, extending the applicability of computational techniques to non-English texts and historical periods. It also will ultimately result in an educational resource as a publicly accessible dataset and toolkits for researchers, educators, and students interested in studying metaphorical language and Francoist literature.

As mentioned before, the dataset that forms the basis of this project requires the compilation of a diverse corpus of novels published during Francoist Spain, considering works from renowned authors as well as lesser-known writers that ensures representation across genres such as historical fiction, romance, and political allegory. For the purposes of this project, I would manually annotate the six selected works for their metaphors. I would include both sentence-level metaphors (often denoted by connecting words and phrases such as ‘as,’ ‘like,’ ‘as though,’ etc.) as well as whole-text allegorical elements found in passages such as the example below from Maute’s *School of the Sun*:

The orchard door, burned by the salt and the wind, was always open (just the opposite of our house, where everything was obstinately locked, as if hidden, as if jealously keeping to the shade.⁶

Throughout the novel, there is a repetitive and significant mention of the house being unlocked, and it is symbolically closed towards the end.⁷ While it is impossible to be objective in this annotation exercise, multiple researchers (representing numerous disciplines, demographics, etc.) working together may form a reasonable though flawed consensus. In a comparable project where researchers attempted to algorithmically detect whether online content was sarcastic, they

⁶ Matute, *School of the Sun*, 130.

⁷ Gonzalez, “Literature of Protest,” 61.

rationalized their collection the following way: “The question we are looking to answer is: *Is the image necessary to find the post sarcastic?* To answer that, we first identify the posts whose sarcastic nature can be positively determined by just looking at the text.”⁸ Similarly, we might ask: *Is reading this text as metaphorical language essential to understanding it?* The corpus will be compiled for the frequency of metaphors used as well as the repetitiveness of certain words. The themes identified by literary scholars working on this topic prior are alienation, isolation, and violence, and data with language that corresponds to these themes will also be collected as well. Though I cannot say without going ahead with the project, I hope to classify the data either by themes such as these or by ‘type’ (allegory, simile, etc., or line-level text versus whole-text metaphors, for example). Some researchers focussed on a singular ‘type.’⁹ This approach can be considered after the data is compiled and cleaned.

Having compiled a corpus of novels published between 1939 and 1975 that represents various genres and authors, the project will then develop a metaphor detection algorithm tailored to analyze Spanish texts from the Francoist period that will be applied to the corpus to identify and categorize metaphorical expressions. Following Tanasescu et al.’s methods, the first step would be to apply rule-based methods to the dataset. These rules mean “[getting] all hypernyms of head nouns and [checking] for each parent till [they] reached the hypernym “abstract entity” or “physical entity,”¹⁰ a method developed by Turney et. al in 2003 to identify literal and metaphorical meanings in both concrete and abstract contexts.¹¹ Turney used “cognitive linguistics that views metaphor as a method for transferring knowledge from a familiar,

⁸ Schifanella et al., “Detecting Sarcasm in Multimodal Social Platforms,” 4.

⁹ Tanasescu, Kesarwani, and Inkpen, “Metaphor Detection by Deep Learning and the Place of Poetic Metaphor in Digital Humanities,” 2.

¹⁰ Tanasescu, Kesarwani, and Inkpen, “Metaphor Detection by Deep Learning and the Place of Poetic Metaphor in Digital Humanities,” 4.

¹¹ Turney et al., “Literal and Metaphorical Sense Identification through Concrete and Abstract Context,” 680.

well-understood, or concrete domain to an unfamiliar, less understood, or more abstract domain,” which attempted to broaden the computer’s semantic understanding of a word.¹² Tanasescu also used a feature based on ConceptNet,¹³ a semantic network that “provides associations between the specific word and real-world knowledge.”¹⁴ Once the dataset has been processed in this way, the project will use a program based on Tanaescu’s freely available web application, which was created with the researchers’ “best performing machine learning model (SVN with F-score 0.781).”¹⁵

The program performs multi-line metaphor detection of a given text using Stanford NLP parser dependencies “to extract the potential word pairs,”¹⁶ though it will need to be recreated to adapt existing metaphor detection algorithms to be suitable for analyzing Spanish texts from the Francoist era. Because the project was initially done in the English language, this adaptation will involve incorporating linguistic features specific to Spanish metaphorical expressions and cultural nuances of the period. After the developed algorithm is applied to the entire corpus to automatically identify and categorize metaphors, the researchers will conduct qualitative analyses to interpret the metaphorical discourse within the socio-political context of Francoist Spain. For example, an idea that is expressed across these texts is that of “internal exile,” a state of being that Gonzalez explains arose from the ways in which “while many intellectuals and artists were forced, or chose, to leave Spain, many remained in their homeland. They existed and produced their art in the political vacuum of Francoist Spain.”¹⁷ A phrase such as this one may

¹² Turney et al., “Literal and Metaphorical Sense Identification through Concrete and Abstract Context,” 680.

¹³ “ConceptNet.”

¹⁴ Tanasescu, Kesarwani, and Inkpen, “Metaphor Detection by Deep Learning and the Place of Poetic Metaphor in Digital Humanities,” 4.

¹⁵ Kesarwani et al., “Metaphor Detection in a Poetry Corpus,” 125.

¹⁶ Kesarwani et al., “Metaphor Detection in a Poetry Corpus,” 125.

¹⁷ Gonzalez, “Literature of Protest,” 18.

need to be clarified such that the program does not assess it as literal. Findings will be compared to Gonzales and other researchers' analysis of the identified metaphors to uncover themes, motifs, and ideological narratives prevalent in Francoist literature. Finally, this work will be examined in the context of the Francoist regime, and the ways in which metaphorical discourse evolved over this duration will be examined considering changes in censorship policies and socio-political contexts.

In Tanaescu's work, it was mentioned in the conclusion that a common point of error in the researchers' analysis was when multiple word senses were not taken into consideration, which skewed the results. Turner writes: "Metaphor is ubiquitous, yet recognizing textual entailment is a challenge when words are used metaphorically. An algorithm for distinguishing metaphorical and literal senses of a word will facilitate correct textual inference, which will improve the many NLP applications that depend on textual inference."¹⁸ Given the nature of the coded language this project aims to identify and decipher, and in a more general sense the elusive and subjective nature of metaphors, this is likely to be a major issue in the accuracy of the results that will arise from this model. Tanaescu's future work "will study a generic concept, trying to track down metaphor based on vector disparity between words and phrases, irrespective of the POS or syntax involved."¹⁹ This is an ambitious goal, and one that may never achieve a 'seamless' accuracy. An additional potential challenge is in the translation of this English-based algorithm into Spanish.

¹⁸ Turney et al., "Literal and Metaphorical Sense Identification through Concrete and Abstract Context," 689.

¹⁹ Tanasescu, Kesarwani, and Inkpen, "Metaphor Detection by Deep Learning and the Place of Poetic Metaphor in Digital Humanities," 8.

That being said, many of the limitations of this dataset pertain to social factors as well. Writing about political cartoons in the digital age, David Francis Taylor explains, “Political cartoons appeal to affluent and highly educated readerships— a kind likely to know something both of literature and of political cartoons past. And they gratify those readers who can register such allusions by reaffirming their cherished sense of cultural capital.”²⁰ This quote makes political cartoons and resistance novels analogous to each other in the sense that the chosen novels may not be a genuine representation of the beliefs of the general public under Franco, and rather may have only pertained to an elite or otherwise select group of people.

One way in which this project may be disseminated effectively is on a website such as the A Homer Commentary in Progress project, which is a “digital commentary supporting the argument that Homer’s language stemmed from a formulaic system of oral poetry.”²¹ The format of the site is accessible, and shows individual pages with highlighted annotations. A page that contains a data visualization of the broader trends that have been found would also be an effective presentation of the metaphor detection. These findings could also be made more accessible to the public via online data visualizations of the most common modes of “code” in the metaphorical language of these works, which would also be an incredibly useful tool in understanding what unifying ideas characterized the writers’ beliefs and criticisms. Since many, if not all of the novels have been made available online on open source platforms, this project is unlikely to be inhibited by copyright, and it would likely not require much sustained upkeep given the limited number of works that it deals with and its historical nature.

²⁰ Francis Taylor, “Political Cartoons in the Digital Age,” 1.

²¹ “Classical Commentaries, The Center for Hellenic Studies.”

In its smallest scope, this project would enrich the academic conversation surrounding resistance literature under Franco — further research could employ distance reading to assess the sheer volume of resistance works that were produced in this period — as of my knowledge, there are currently no statistics that give us a sense of this. The work would contribute to a more general understanding of how individuals find ways to express themselves under repressive regimes. This research proposal outlines a comprehensive investigation into metaphorical discourses in novels from Francoist Spain, employing digital humanities methodologies to uncover hidden meanings and ideological subtexts. By analyzing metaphor usage across a diverse corpus of texts, this project aims to contribute to both literary scholarship and historical understanding of the Francoist period, offering insights into the complexities of cultural production under authoritarian regimes.

Works cited

1. “Classical Commentaries, The Center for Hellenic Studies.” Accessed May 3, 2024. <https://ahcip.chs.harvard.edu/about-1>.
2. “ConceptNet.” Accessed May 3, 2024. <https://conceptnet.io/>.
3. Francis Taylor, David. “Political Cartoons in the Digital Age.” *Yale University Press* (blog), June 27, 2018. <https://yalebooks.yale.edu/2018/06/27/political-cartoons-in-the-digital-age/>.
4. Gonzalez, Margaret. “Literature of Protest: The Franco Years.” *LSU Historical Dissertations and Theses*, January 1, 1994. https://doi.org/10.31390/gradschool_disstheses.5726.
5. Kesarwani, Vaibhav, Diana Inkpen, Stan Szpakowicz, and Chris Tanasescu. “Metaphor Detection in a Poetry Corpus.” In *Proceedings of the Joint SIGHUM Workshop on Computational Linguistics for Cultural Heritage, Social Sciences, Humanities and Literature*, 1–9. Vancouver, Canada: Association for Computational Linguistics, 2017. <https://doi.org/10.18653/v1/W17-2201>.

6. Matute, Ana María. *School of the Sun*. Columbia University Press Morningside ed. Twentieth-Century Continental Fiction Series. New York: Columbia University Press, 1989.
7. Perez, Janet. "Fascist Models and Literary Subversion: Two Fictional Modes in Postwar Spain." *South Central Review* 6, no. 2 (1989): 73. <https://doi.org/10.2307/3189557>.
8. Schifanella, Rossano, Paloma de Juan, Joel Tetreault, and Liangliang Cao. "Detecting Sarcasm in Multimodal Social Platforms," 2016. <https://doi.org/10.48550/ARXIV.1608.02289>.
9. Tanasescu, Chris, Vaibhav Kesarwani, and Diana Inkpen. "Metaphor Detection by Deep Learning and the Place of Poetic Metaphor in Digital Humanities." *The Thirty-First International Florida Artificial Intelligence Research Society Conference (FLAIRS-31)*, n.d.
10. Turney, Peter, Yair Neuman, Dan Assaf, and Yohai Cohen. "Literal and Metaphorical Sense Identification through Concrete and Abstract Context." In *Proceedings of the 2011 Conference on Empirical Methods in Natural Language Processing*, edited by Regina Barzilay and Mark Johnson, 680–90. Edinburgh, Scotland, UK.: Association for Computational Linguistics, 2011. <https://aclanthology.org/D11-1063>.