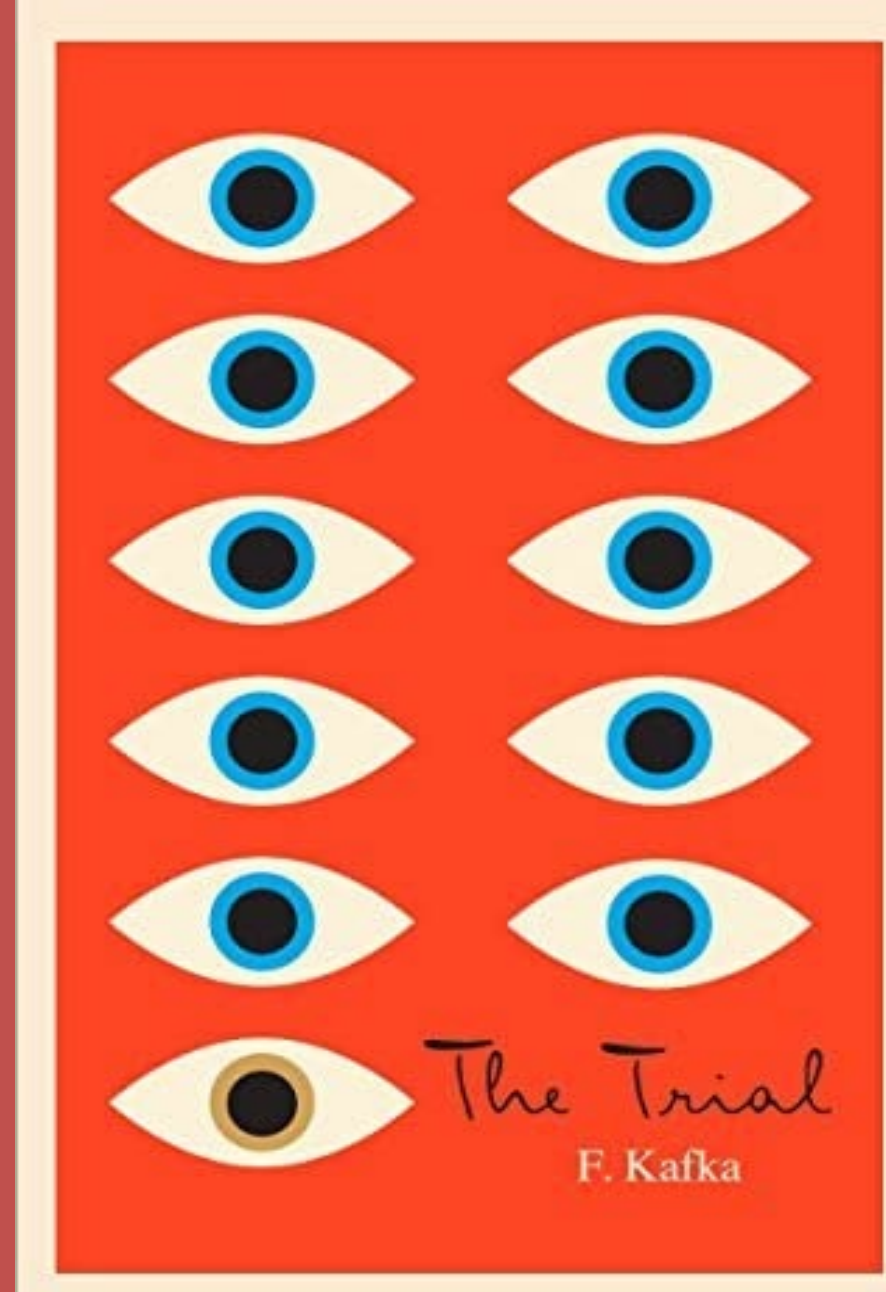




The Trials of Translation: A Cross -Linguistic Survey of Sentiment Analysis on Franz Kafka's Flannery Strain

IPHS 200 Programming Humanity (Fall 2022) Prof Elkins and Chun, Kenyon College

Trial



Abstract

“A novel,” wrote Umberto Eco in *The Name of The Rose*, “is a machine for generating interpretations.”

The practice of translation and sentiment analysis share a similarly pressurized focus on the interpretation of the meaning and effect of a small byte of language within its narrative context. In these interpretations though, a fascinating point of divergence arises – whereas sentiment analysis software follows an algorithmic sentiment scoring when determining the significance of a word, human translators must grapple with how to rightfully bring that word into another language. The German-speaking Czech writer Franz Kafka is a writer who has baffled translators and literary critics alike since his initial publication, many who have despaired and rejoiced at the seeming impossibility of translating parts of his works.

Curious about how the myriad translations of one of Kafka’s most known works, *The Trial* (1925), sit in narrative position to one another, I used a natural language processing model to analyze the sentiment arc among the English, Spanish, French, and German translations. The question here was: Do these sentiment analysis tools work multi-linguistically, and if so, are the crux moments of significance that surface throughout the analyses different across languages?

Introduction

The Trial, written 1914-15, is an unfinished novel of Kafka’s, published posthumously in 1925. The novel follows the character of Joseph K. through a mysterious legal system, as he is accused of an unknown crime for no apparent reason, and becomes powerlessly tangled in the trappings of the court, until he is eventually executed, still without awareness of the circumstances of his fate.

In his book *Transforming Kafka: Textual Effects*, Patrick O’Neill explores the creation of a “worldwide Kafka system,” taking a macro-textual approach to survey the effects of translation in Kafka’s work. Known for his bizarre stories featuring themes of alienation, anxiety, and bureaucratic absurdity, Kafka’s writing and literary construction is so distinctly nightmarish that the term “Kafkaesque” has emerged as an eponymous adjective to describe sensations of dangerous, surreal, disorientation. It is particularly befuddling because of its indeterminate and destabilized narrative focalization – it is hard to tell often if a statement is intended to be understood as coming from a character, or from the narrator of the story. O’Neill finds an example of this empirical uncertainty in the very first sentence of *The Trial* – “Someone must have been telling lies about Joseph K...” – a sentence that has been a subject of contention for years amongst translators and editors, with varying levels of ambiguity applied to “someone,” varying verbs to describe “telling lies” (alternates - *slandered*, *defamed*, *made a false accusation*). The second part of the sentence has a slew of different translations as well: “...for without having done anything wrong he was arrested one fine morning” (as found in 7 of 9 English versions), also interpreted as “he knew he had done nothing wrong” (Wyllie’s English version). Within the first sentence of the novel alone, we can see the diverse forms that word choices construct in meaning.

Kafka’s oeuvre has been translated into more than 40 languages to date, but for this project I focused only on Spanish, French, English, and *The Trial*’s original German. Though Scottish couple Willa and Edwin Muir have the preeminent English translators of Kafka, for our English model, I used David Wyllie’s translation, the version available in the public domain through Project Gutenberg. Wyllie’s translation was published in 2007, in the wake of Sir Malcolm Pasley’s critical edition. The French translation I used was Alexandre Vialatte’s 1933 *Le Procès*, many times revised and republished by Gallimard in 2006. The Spanish translation was published in 2014 by Edif Editions, proceeding Vicente Mendivil’s original translation of *El Proceso* in 1939. The German text used, was of course, not a translation but the original text of *Der Prozess*, published in 1925 in Berlin by Verlag Die Schmiede.

Methodology

With the help of Professor Chun and Professor Elkins, we used a bert-base-multilingual-uncased model fine tuned for sentiment analysis. The model is intended for use in six languages: English, Dutch, German, French, Spanish and Italian. It is modeled for sentiment analysis of product reviews, but for our purposes, I used it for Kafka’s novella. The NLP model was fine tuned through analysis of 5,000 product reviews in various languages, though some languages received a bit more fine tuning than others. The accuracy of the model for each language is reflected in its prediction of the number of stars a human reviewer would give based on the review: Accuracy (Exact): English 67%, German 61%, French 59%, Spanish 58%. Accuracy (Off-by-1): English 95%, German 94 %, French 94%, Spanish 95%. Given these numbers, I feel confident that the model has done a reasonably accurate job in following the sentiment arcs of the four versions of *The Trial*.

I also ran the English text through Professor Chun’s multi-model sentiment analysis code that uses Vader, TextBlob, DistilBERT, MultiBERT, and Fine-Tuned Sentiment RoBERTa.

Once the cruxes were plotted, Professor Elkins and I analyzed the various valleys and peaks across the four languages, looking for major similarities and differences in interpretations of narrative moments. As I am not fluent in French, German, or Spanish, I had to use Google Translate to conceptualize the corresponding text within the entirety of the story.

Results

The graphs returned to us by the sentiment analysis model at first appear pretty dissimilar in terms of superficial shape. Upon closer examination and connection to textual context, there are revealing connections in the points that were highlighted.

The novel begins with a downward turn across the four texts. The first English “valley” includes K’s conversation with Frau Grubach, in which he assures her that the disturbance caused by his arrest that morning “will not happen again.” (EN 333). Here, K is withholding, and what follows is his solitary descent into the legal system about which he struggles to communicate. The German, English, and French models all show a significant early valley at the moment when K is informed by Fraulein Montag that his fellow lodger who he is interested in, Fraulein Burstner, is disinterested in him. (FR 1160, GR 912, EN 1292). The initial most notable valley in the Spanish model is K’s approach to dusty courts in the attic, when he is overcome by bouts of dizziness (SP 1156).

All four models display an upward trajectory as K’s optimism builds on previous despair and confusion with hope of K’s uncle’s friend the lawyer, and the eventual meeting of the alluring Leni (climaxes at FR 1555, GR 1260, EN 1610, SP 1585). A swift disillusionment with the lawyer’s capability and any hope for K’s absolution is seen in the sudden decline that follows across the four models. A second point of unanimity is the instance when K is introduced by the manufacturer to Titorelli, a painter from whom K seeks out advice about his legal situation, though he leaves even more disheartened (FR 2375, EN 2302, SP 2297, GR 1863). Another pronounced unanimous valley is K’s meeting with a visiting Italian businessman, but the four models swing in wildly different directions before reaching the “Before the Law” parable. The ending of the novel also resists a unanimous interpretation across the models.

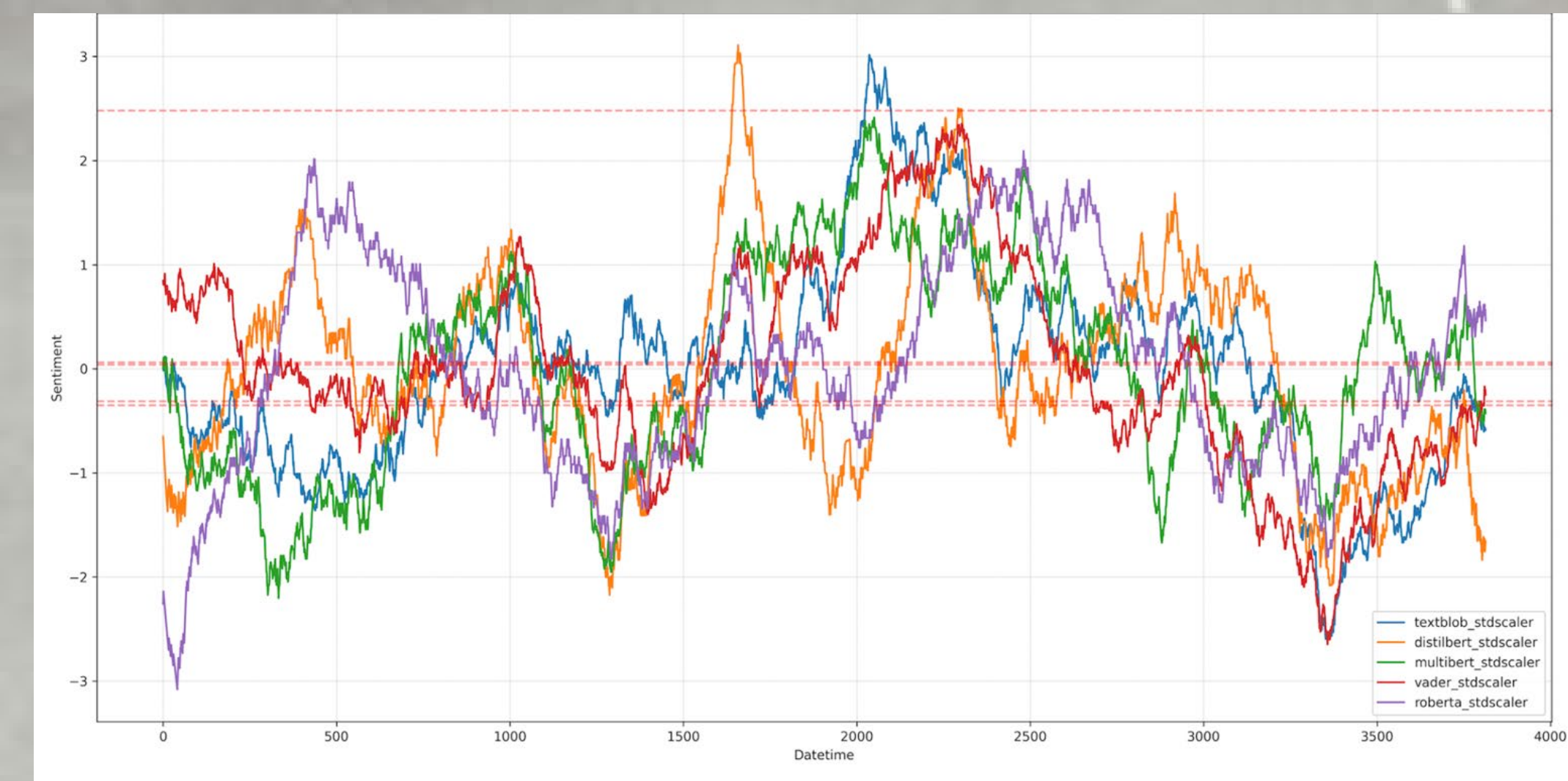
Conclusion

While the models show marked differences, an attendance to those differences reveals that most of the plot points, dialogue excerpts, and character’s actions, that were spotlighted as significant found similarities among the models. While I’d expected the sentiment analysis program to be most accurate towards English text and least towards Spanish (based on its experience in volume of product reviews), the analysis performed comparably well with the Spanish model as it did with the other three. After completing the analysis, I realized that the French version had an epilogue, which may also account for graphical differences.

On the whole, it seems that the models did fairly well in predicting the gravity of scenes for the larger narrative – though layers of nuance could be argued for each moment of significance in Kafka’s writing. This is the level of analysis undertaken by O’Neil in *Transforming Kafka*, that goes above and beyond and much more specific than my cross-textual analysis here.

It must also be remembered that the novel is unfinished, and that there was also a chapter deleted. *The Trial* is a tricky novel to parse, and graphical renderings of sentiment arcs by human readers of *The Trial* would likely not reflect one another at all, which I believe is partly why the models vary so much. This impermeable ambiguity of Kafka’s prose is further seen in the multi-model English graph (*below*), that displays relatively wide variation.

MULTI-MODEL ENGLISH



Future and Ethics Statement

This project would have been helped exponentially if I had beyond an elementary comprehension of Spanish, and any knowledge of French or German. It seems that my use of Google Translate does a great dishonor to those distinguished translations, and not to mention to the original text. If given more time, I would have liked to have dived deeper into the English text and thought more critically about the crux points, how characters reappear amongst them, and the particular language that drives those analysis choices.

In any case, I look forward to seeing the development of more cross-linguistic sentiment analysis models, as well as the development of the field of Kafka translation.

References/Acknowledgements

Thank you to Professor Elkins and Professor Chun for their tireless guidance and support throughout this project.

Lemon, Robert (2016) “Patrick O’Neill, Transforming Kafka: Translation Effects. Toronto: U of Toronto P, 2014. 222 pp.” *Studies in 20th & 21st Century Literature*: Vol. 40: Iss. 1, Article 8. <https://doi.org/10.4148/2334-4415.1368>

Kafka, Franz. *Der Prozess*. Verlag Die Schmiede, 1925. Online Text courtesy of Project Gutenberg.

Kafka, Franz. *El Proceso*. Biblioteca Edaf, 2014.

Kafka, Franz. *Le Procès*. Editions Gallimard, 2006.

Kafka, and Wyllie. “The Trial by Franz Kafka.” *Project Gutenberg*, 1 Apr. 2005. <https://www.gutenberg.org/ebooks/7849>.