Cultural Analytics

ENGL 64.05

Fall 2019

Prof. James E. Dobson

October 30, 2019

The Affective Fallacy

- W. K. Wimsatt, JR., and M. C. Beardsley published an essay title "The Affective Fallacy" in 1949.
- It was built on prior arguments they had made in an essay titled "The Intentional Fallacy" (1947).
- They seek to re-center the text as the object of analysis:
  - "The Affective Fallacy is a confusion between the poem and its *results* (what it *is* and what it *does*)" (31).
  - "The outcome of either Fallacy, the Intentional or the Affective, is that the poem itself, as an object of specifically critical judgment, tends to disappear" (31).
- These essays helped, in part, legitimize an "objective" literary criticism by removing subjective feeling about the work and biographical facts and information originating in the author.

DARTMOUTH
1769
VOX CLAMANTIS IN DESERTO

Sentiment Mining

Sample Sentences of a Subjective Review

"(1) I bought an iPhone a few days ago. (2) It was such a nice phone. (3) The touch screen was really cool. (4) The voice quality was clear too. (5) Although the battery life was not long, that is ok for me. (6) However, my mother was mad with me as I did not tell her before I bought it. (7) She also thought the phone was too expensive, and wanted me to return it to the shop. … "

Bing Liu: "Sentiment analysis or opinion mining is the computational study of opinions, sentiments and emotions expressed in text" (Liu 2010).

DARTMOUTH 1769 VOX CLAMANTIS IN DESERTO

Key Concept: Opinion Lexicon

"In the research literature, opinion words are also known as *polar words*, *opinion-bearing words*, and *sentiment words*. Positive opinion words are used to express desired states while negative opinion words are used to express undesired states. Examples of positive opinion words are: *beautiful*, *wonderful*, *good*, and *amazing*. Examples of negative opinion words are *bad*, *poor*, and *terrible*. Apart from individual words, there are also opinion phrases and idioms, e.g., *cost someone an arm and a leg*. Collectively, they are called the *opinion lexicon*. They are instrumental for sentiment analysis for obvious reasons" (Liu 2010).

Base Type / Comparative Type

- Opinion words can be categorized as base types. There are other types of language involved in subjective opinions.
- These can be comparative:
  - better than, less than.
- Or superlative:
  - best, worst, greatest

DARTMOUTH
VOX CLAMANTIS IN DESERTO
1769

Key Concept:
Text-External Referential Systems

- Data that are external to the world of modeled texts that are used in the processing of make meaning from extracted data within the model.
- I argue that "secondary databases within computational criticism function, in perhaps several senses, as heuristic aids for making meaning of the analyzed texts. Secondary databases do not rise to the same level, to the primacy as an object of attention or critique, as the major objects under examination because secondary databases are imagined as functioning only to assist in the interpretation of another digital object. The primary use of secondary sources at present is to assist in locating objects in space and in time, even if users do not recognize that this is how they are using them" (84).
- I write that the "rapidly expanding use of secondary digital sources, data sources that because of their nature or idiosyncratic structure fall outside the carefully edited and annotated textual apparatus, are also removed from any detailed attempt to record bibliographic metadata. These secondary objects—for example, databases, dictionaries, statistical models, and maps—are potential sources of knowledge about the world, but insofar as they are cultural objects, they cannot be treated as ahistorical repositories of truth" (79-80).

DARTMOUTH 1769 VOX CLAMANTIS IN DESERTO

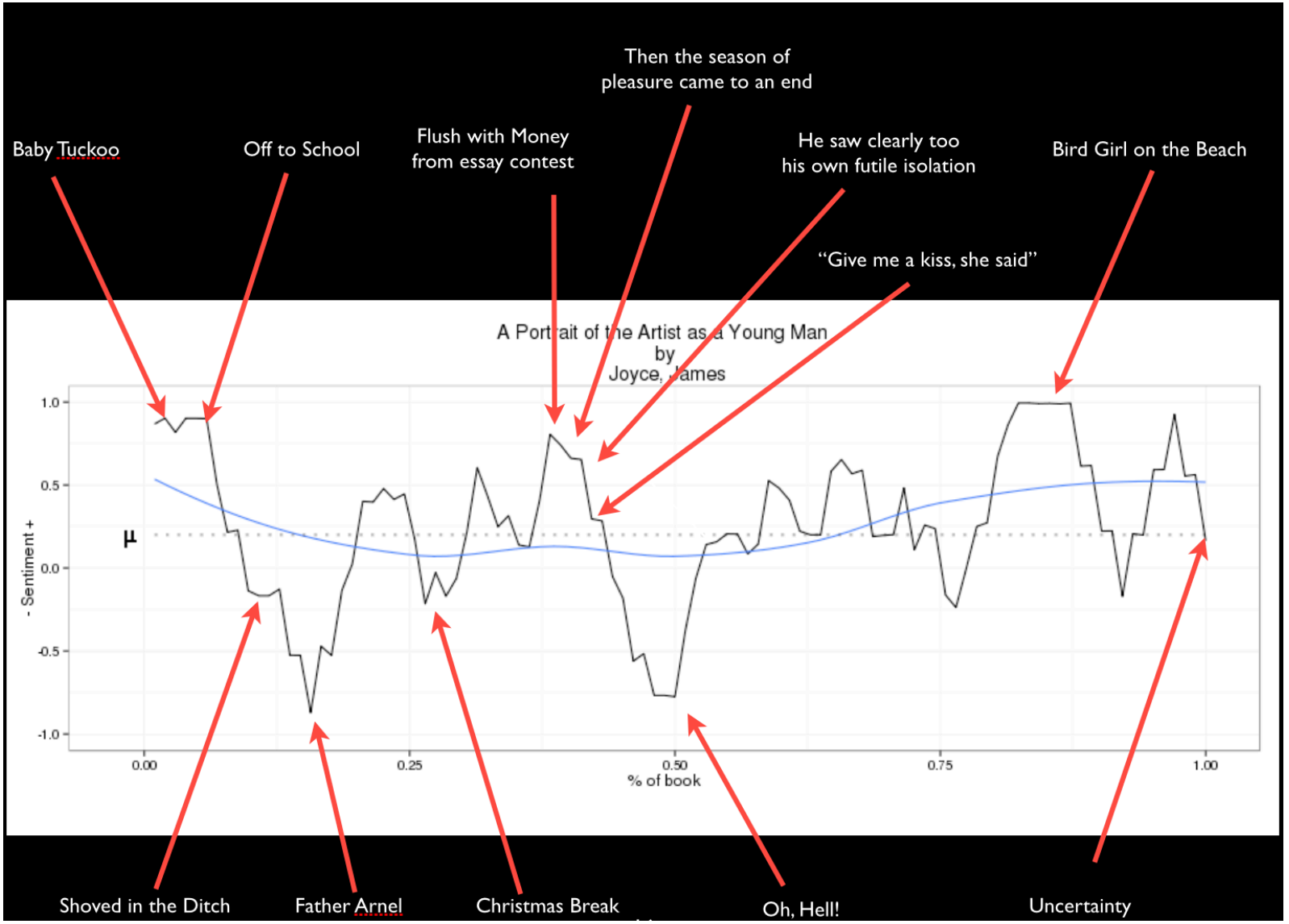
Sentiment as Proxy for Plot

- Matthew Jockers published a series of blog posts, beginning in 2014* on using sentiment analysis.
- He writes that "by accident I discovered that the sentiment I was detecting and measuring in the fiction could be used as a highly accurate proxy for plot movement."
- Jockers created a package for R to examine sentiment in fiction called syuzhet. He gave it this name after the Russian formalists to name the plot as presented to readers: *"Syuzhet* is concerned with the linear progression of narrative from beginning (first page) to the end (last page), whereas fabula is concerned with the specific events of a story, events which may or may not be related in chronological order. When we study *fabula*, which is what we typically do in literature courses, we mentally reconstruct the events into chronological order. We hope that this reconstruction of the *fabula* will help us understand the experience of the characters, the core story, etc. When we study the *syuzhet*, we are not so much concerned with the order of the fictional events but specifically interested in the manner in which the author presents those events to readers."

* http://www.matthewjockers.net/2014/06/05/a-novel-method-for-detecting-plot/

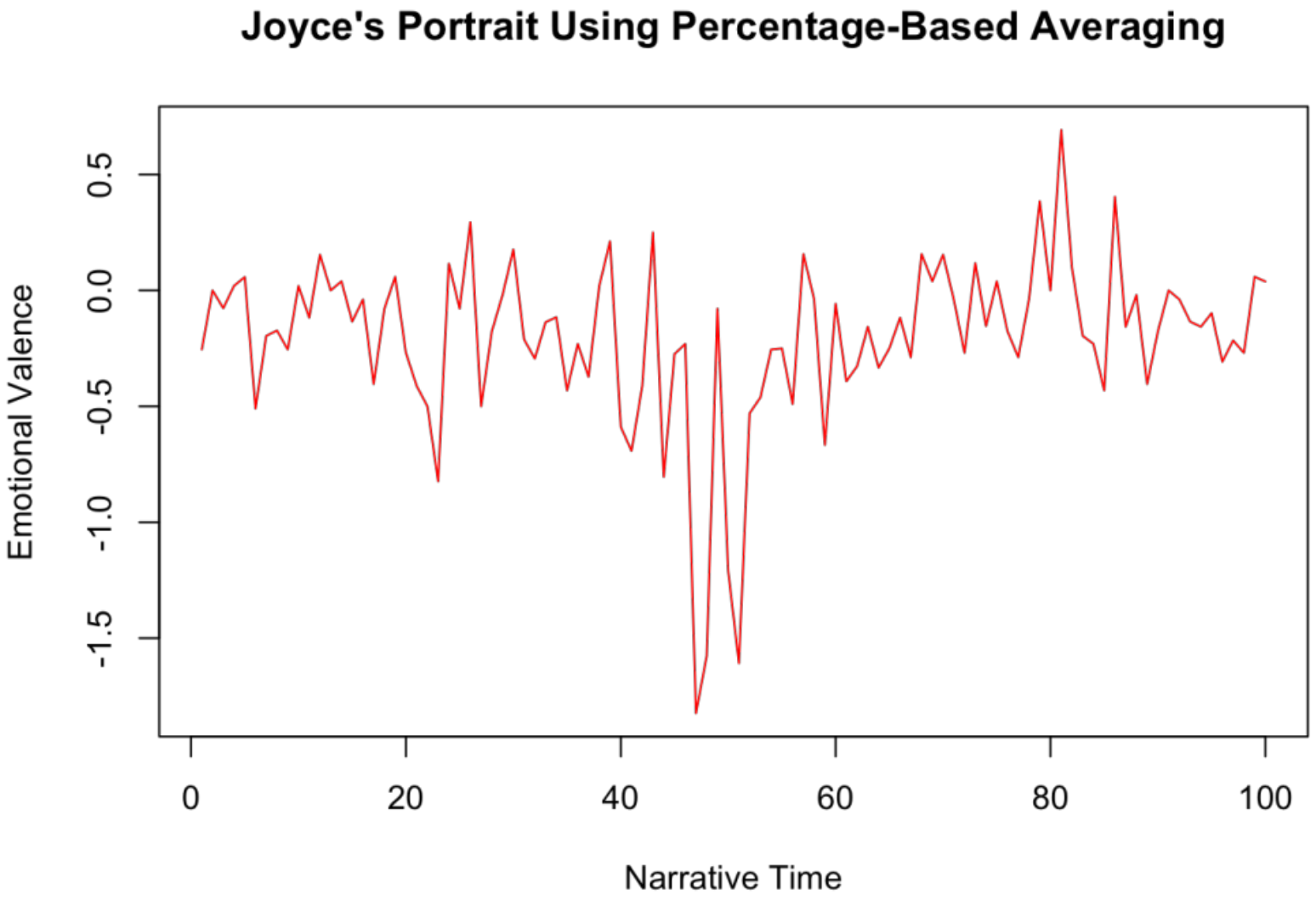
DARTMOUTH 1769 VOX CLAMANTIS IN DESERTO

Cultural Analytics



From Jockers's website

DARTMOUTH
1769
VOX CLAMANTIS IN DESERTO



From Jockers's website

DARTMOUTH
1769
VOX CLAMANTIS IN DESERTO

"The Emotional Arcs of Stories are Dominated by Six Basic Shapes"

Reagan et al.

Expressly interested in the emotional arc, not the plot:

> "While the plot captures the mechanics of a narrative and the structure encodes their delivery, in the present work we examine the emotional arc that is invoked through the words used. The emotional arc of a story does not give us direct information about the plot or the intended meaning of the story, but rather exists as part of the whole narrative (e.g., an emotional arc showing a fall in sentiment throughout a story may arise from very different plot and structure combinations). This distinction between the emotional arc and the plot of a story is one point of misunderstanding in other work that has drawn criticism from the digital humanities community" (2).

DARTMOUTH 1769 VOX CLAMANTIS IN DESERTO

Cultural Analytics

projects ▾ About Instructions Words Blog Press Papers Talks Fund us Data Consulting

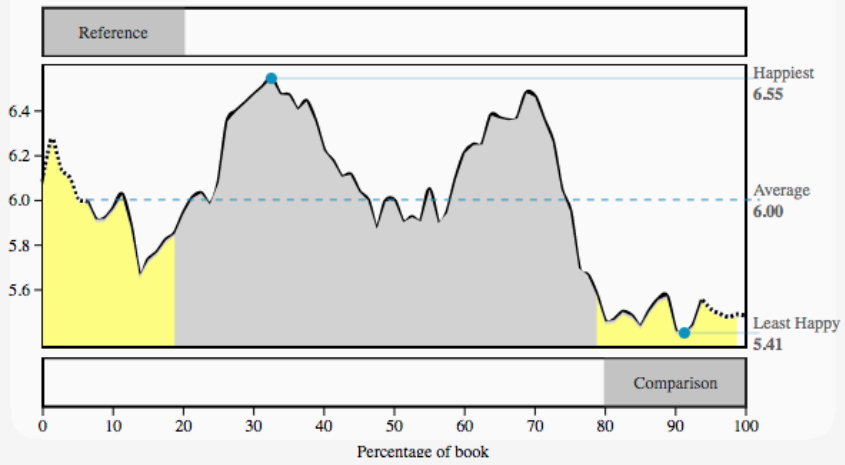
Harry Potter and the Sorcerer's Stone

by J.K. Rowling

Search Books | by Title ▾ | Classics ▾ | Harry Potter ▾ | Random

Book happiness time series:

Explore the work's emotional dynamics by sliding and resizing the reference and comparison sections.

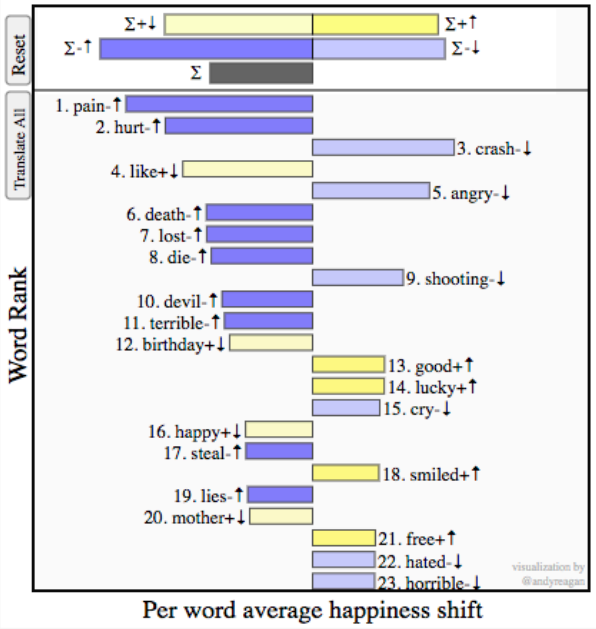


Word Shift:

Why comparison section is less happy than reference section:

Reference section's happiness: 5.87

Comparison section's happiness: 5.50

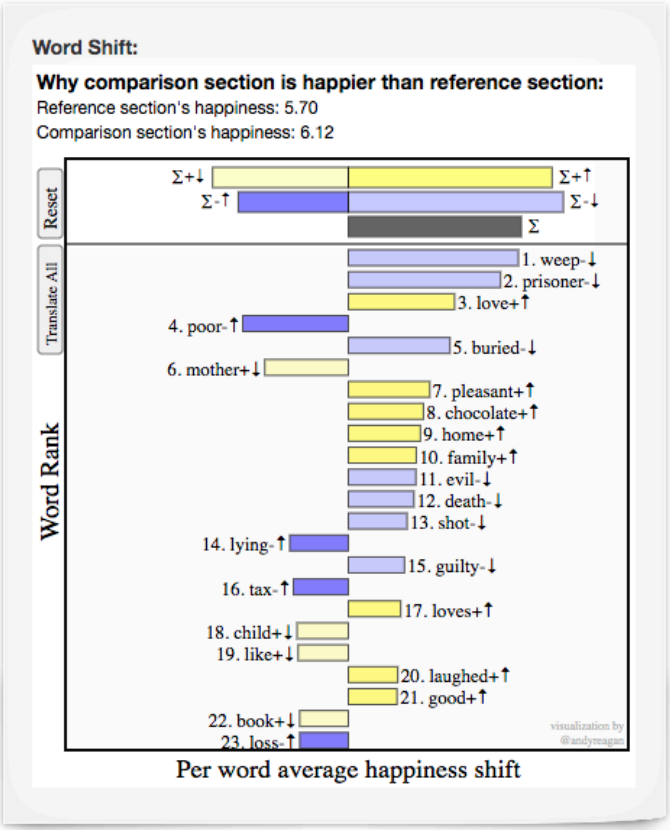


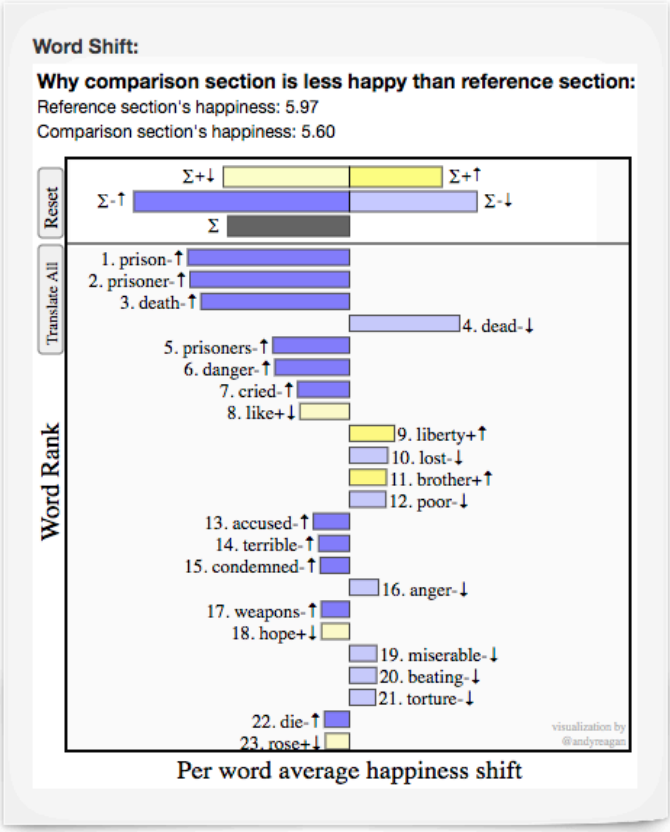
Lens (for advanced users):

Slide and resize the stop-window to change the lens:

DARTMOUTH 1769 VOX CLAMANTIS IN DESERTO

Cultural Analytics

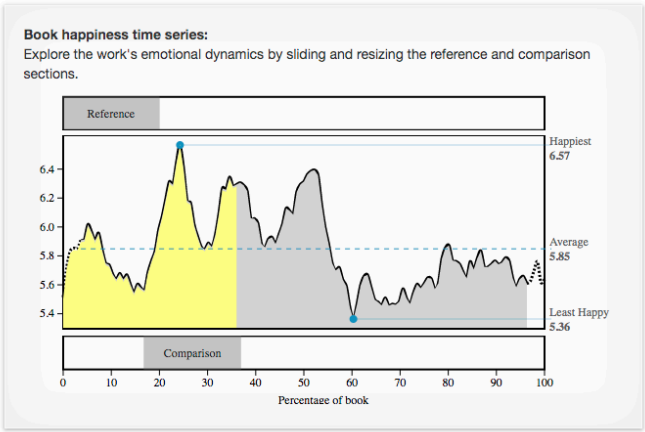




A Tale of Two Cities (wiki)

Search Books

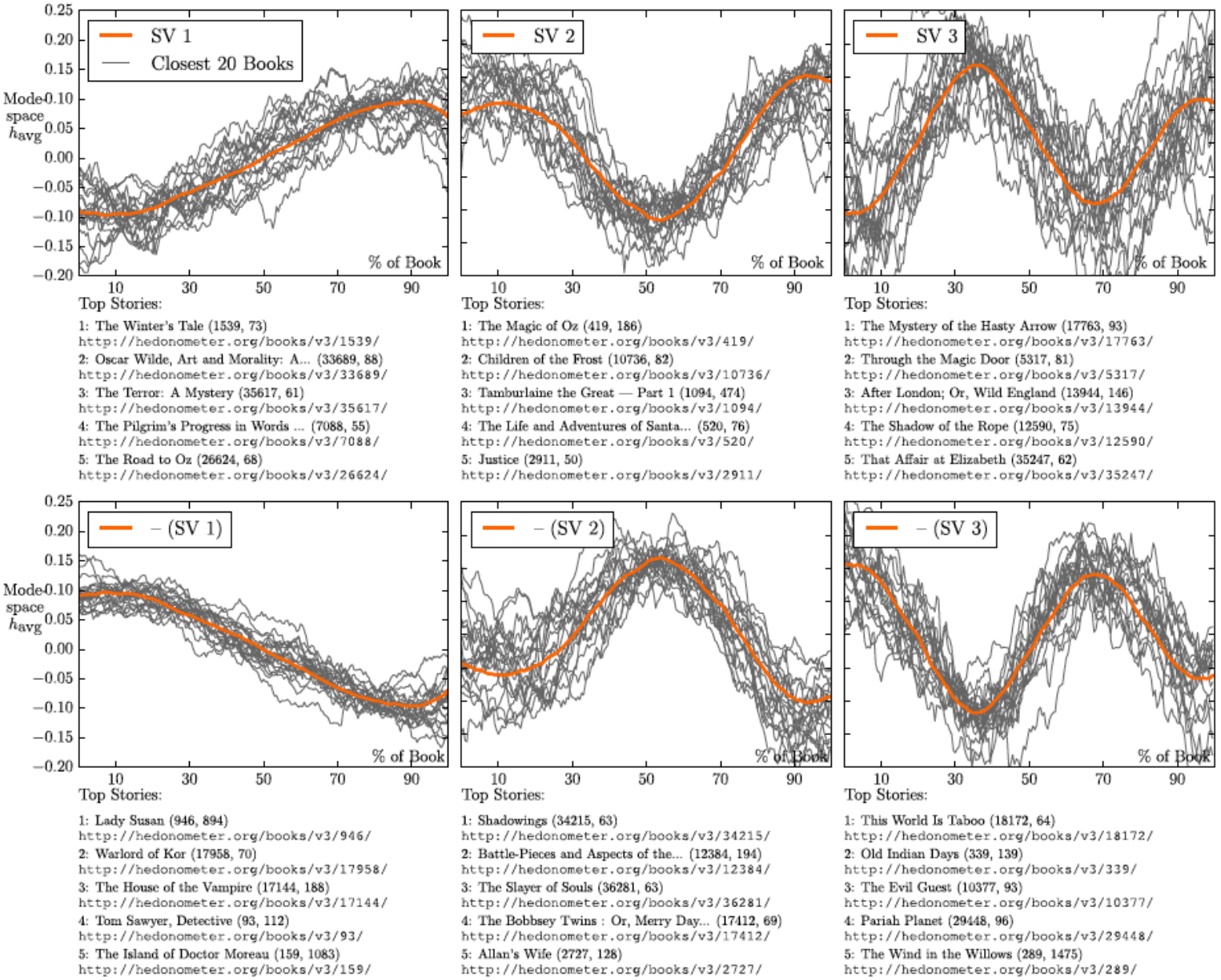
by Charles Dickens



DARTMOUTH
1769
VOX CLAMANTIS IN DESERTO

Jockers on emotional markers and plot proxies:

"The Suyzhet package performs a certain type of text analysis, and I'm claiming that the results of this analysis may serve as a pretty darn good proxy for plot. That said, I've been working on this problem for two years, and I know some specific places where it fails. The most spectacular example of failure was discovered by my son. He'd just finished reading one of the books in my corpus, and I showed him the plot shape from the book and asked him it it made sense. He said, "well, yes, mostly. But this spike here is all wrong." It was a spike in good fortune, positive valence, at precisely the place in the novel where the villains had scored a major victory. The positive valence was associated with a several page long section in which the bad guys were having a very good time. Readers, of course, would see this as a negative moment in the text, Suyzhet does not. Nor does Suyzhet understand irony and dark humor and so on. On a whole, however, Suyzhet gets it right, and that's because *most* books are not sustained satire, or sustained irony. Most books end up using emotional markers in a fairly consistent and conventional way. Indeed, even for an experimental novel such as Joyce's *Ulysses*, Suyzhet produces a plot shape that I consider to be a good match to the ebbs and flows of the text."



Closest SVs (Reagan et al)

"We can immediately recognize the familiar shapes of core emotional arcs in the first four modes, and compositions of these emotional arcs in modes 5 and 6.We observe 'Rags to riches' (mode 1 positive), 'Tragedy' or 'Riches to rags' (mode 1, negative), Vonnegut's 'Man in a hole' (mode 2, positive), 'Icarus' (mode 2, negative), 'Cinderella' (mode 3, positive), 'Oedipus' (mode 3, negative). We choose to include modes 7-12 only for completeness, as these high frequency modes have little contribution to variance and do not align with core emotional arc archetypes from other methods" (Reagan et al, 5).

DARTMOUTH
1769
VOX CLAMANTIS IN DESERTO