



# Detail in the Novel

## Tim Schott

School of Information, MIMS '23  
University of California, Berkeley

### ABSTRACT

Outside of eloquently-written conjecture, literary scholars lack a clear definition of fictional details and their implications – when it comes to particulars, they’re anything but (Miles, 1979; Auerbach, 1946). The objective of this project is to **analyze a corpus of English-language novels and evaluate their usage of details**. I adopt a mixed methods approach, spanning the methodological gamut from frequency-based analysis to fine-tuning and interrogating large language models.

### DATA

I randomly sampled **13 passages of approximately 800 characters from 28 canonical novels** with a “reputation” for realism, exacting description, etc. such as Émile Zola’s *Germinale*. I **annotated their usage of detail**. 197 detail and 167 not\_detail samples comprise baseline dataset for experiments.

### METHODS

I pursued multiple routes of analysis with this dataset, combining computational and qualitative analysis.

**Specificity:** How can we assess the **granularity** of a sample’s diction? Relying on WordNet, a lexical database that catalogues semantic relationships, we can make judgements about each sample. I calculate a specificity score by averaging each noun and verb’s distance from its broadest hypernym (Nelson 2020).

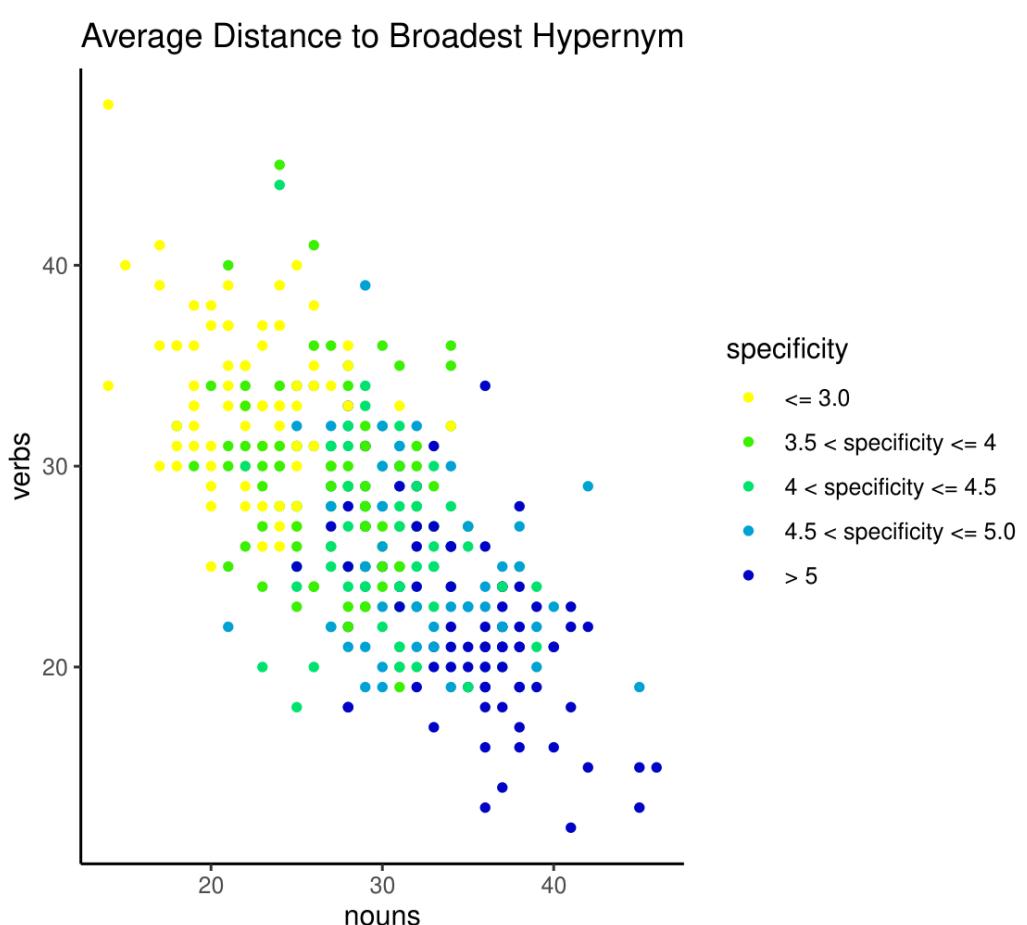
**Preposition Analysis:** Detail in the novel is thought to be serviced by the use of **prepositions** (Genette, 1976). To evaluate whether the detail samples use prepositions **significantly** more than the control set, I performed the Chi-squared test of independence for each preposition (Bamman 2021)

**BERTology:** Which words within my dataset are the most **influential**? I fine-tuned a BERT classifier on my dataset and analyzed the specific words that influenced the model’s classification boundaries (Devlin et al., 2019, Patzke et al., 2019)

### RESULTS

#### Specificity:

- detail passages are **14%** more specific and use **17%** more nouns and adpositions



*Left:* Nouns are more specific than verbs. And the granularity of those nouns matters as well: is a character erranding to the *market* or the (more-specific) *patisserie*?

#### Preposition Analysis:

- Over **half** of prepositions occur **significantly** frequently in detail passages (via Chi-squared).
- 6 common prepositions such as *down*, *in*, and *through* occur **exclusively** in detail passages.

#### Bertology:

- BERT mirrors the sensibilities of critics and human annotators!

Subset of Top 100 most attended to tokens  
via Captum on test data

token	weight	class
at	1.5794992	preposition
like	0.4044355	preposition
across	0.2941255	preposition
against	0.2761270	preposition
of	0.2509571	preposition
by	0.2498414	preposition
behind	0.2076101	preposition
under	0.2012971	preposition
between	0.1765518	preposition
eight	0.1476980	number
ten	0.1353989	number
twice	0.1348596	number
saw	0.1618875	meta
observed	0.1364293	meta

*Left:* Sum of attention weights after fine-tuning BERT-medium (via Captum). Prepositions and numbers are of outsized importance, as are *saw* and *observed*.

*Below:* How BERT reads a passage. **Green** highlights represent words associated with detail, including prepositions like *near*.

### CONCLUSION

The goal of the project is not the production of a state-of-the-art model capable of recognizing what is and is not a detail, rather, it is to **create the conditions** in which constructing such a model is for the first time **a real possibility**.

I produce original insights, highlighting the **significance of nouns and prepositions** in my corpus, and also create computational pipelines that will aid future work in this domain. I present a **novel integration of a large-language model** that corroborates my frequency-based explorations.

### FUTURE WORK

These experiments represent exciting first steps to better understand the function of detail in the novel. Numerous areas relevant to computational humanists are yet to be explored:

**Contemporary Literature:** Incorporate contemporary fiction from *The Contemporary Litbank* (Bamman, 2019).

**Literary Theory:** Why do details “hook” us (Felski, 2020)? Do concepts like mimesis, realism and narratology withstand computational scrutiny (Lukacs 1914, Auerbach, 1946, Barthes and Duisit, 1989)? If novels are “works of art which arrive at truth by telling untruths,” as Tolstoy biographer A.N. Wilson writes, **are details vectors of truth or lies?**

**Quantitative Formalism:** Expand contemporary studies of digital narratology (Wilmot and Keller, 2021) and Russian Formalism (Lvoff, 2011). What objects and people are privileged to warrant detail? Are there patterns to the employment of detail across and within narratives?

**Annotation Guidelines:** Expert annotated datasets are in short supply in the digital humanities. I would like my experiments to help produce **peer-produced detail identification guidelines**.

### ACKNOWLEDGEMENTS

I am very thankful for the School of Information faculty who offered guidance on my project, in particular David Bamman and Coye Cheshire. This project was completed as part of Professor Bamman’s *Applied Natural Language Processing* course (INFO 256, Fall 2021) and will be expanded Fall 2022 through an independent study.

To view my source code as well as references for this presentation, please visit the project Github repository: [github.com/timschott/details](https://github.com/timschott/details)

[CLS] ... go away from the pit , from which he was driven off ; he could not even sit down and wait near his mother , he was so deeply stirred by the need to act , which drove him constantly on . it was the third day . [UNK] in despair , had resolved to abandon the attempt in the evening . at midday , after lunch , when he came back with his men to make one last effort , he was surprised to see zach ari e , red and ge stic ultating , come out of the mine shouting : she ' s there ! she ' s replied to me ! come along , quickly ! he had slid down the ladder s , in spite of the watch man , and was declaring that he had heard hammering over there , in the first passage of the guillaume seam . but we have already been twice in that direction , [UNK] observed , sc ept ically . any how , we ' ll go and see . ma he ude had risen , and had ... [SEP]