Detail in the Novel

Tim Schott
School of Information
University of California, Berkeley
timschott@berkeley.edu

1 Introduction

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, 2003) 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 training and interrogating large language models. Following these preliminary investigations, I plan to collaborate with relevant parties (hackers and humanists alike) and furnish a coherent annotation framework that can be used to identify the level of detail used throughout a corpus.

2 Related Literature

2.1 Narratology, Salience, Details

Literary critic Roland Barthes provides theoretical foundation for the field of narratology when he taxonomizes components of narrative (Barthes and Duisit, 1975). One of the elements he identifies is the "cardinal function." A cardinal function refers to an action that "directly [affects] the continuation of the story . . . it either initiates or resolves an uncertainty." Barthes also narrows his gaze and grapples with details. In addition to cardinal functions, every narrative possesses a certain number of details (Barthes, 1989). He contrives a category called "informants" that produce "ready-made knowledge" and exist "to root fiction in the real world (Barthes and Duisit, 1975). Short of a few examples like providing a character's precise age, Barthes is unfortunately laconic in his explanation of this phenomenon. Regardless, from Barthes' writing we conclude that any method of cataloguing details needs to encode occurrences of familiarizing information. These facets should be readily identifiable with examples such as cardinal directions, numbers, colors, and other material realities. To wit, a close-reading of Leo Tolstoy's realist fiction posits these sort of empirical details are significant because they contribute to a discourse that generates persons and worlds that feel authentic (Auyoung, 2015).

Meanwhile, fellow narratologist Gerard Genette opposes "narration" and "description" (Genette, 1976). The mode of narration represents actions and events while description depicts objects and people (Genette, 1976). Description in prose, according to Genette, "lingers over objects and ... seems to suspend the flow of time... successively [modulating] the representation of objects simultaneously juxtaposed in space" (Genette, 1976). This bifurcation parallels Viktor Shklovsky's dyad of "fabula" (story) and "syuzhet" (plot) where syuzhet serves as "a phenomenon of style" and allows the author to augment a narrative core that can always be reduced to "Because A, Because B, etc." (Liveley, 2019).

Genette's argumentation mirrors Barthes. Specific cases of this conjectured "modulation" occurring with a syuzhet are scant. Lack of examples notwithstanding, Genette offers justification for contending that a novel's descriptive power lies in its capacity to vividly depict and fictional objects (Genette, 1976). Other scholars have attempted to extend this idea, such as advancing a formal conception of a "descriptive detail," but the analyses are likewise plagued by generalities and a lack of reproducible steps for identification (Schor, 1984). Regardless, their work guides the trajectory of both this survey and other digital efforts.

2.2 Digital Narratology

The most germane computational research for this effort operationalizes Barthes' theory of salience. For instance, a fascinating study fuses Barthe's no-

tion of the cardinal function with the **BERT** large language model; the authors take Barthes at his word and operationalizes cardinal functions to determine whether the **BERT** can effectively detect important sentences across a corpus of folktales (Devlin et al., 2019; Otake et al., 2020). This study considers a sentence to be "salient" if its omission from the story greatly reduces the narrative's coherence (Otake et al., 2020). Furthermore, a duo of prolific computational narratologists are producing excellent studies regarding tasks such as identifying suspense (Wilmot and Keller, 2020). The pair enhances the technique used by Otake et al. (2020) with a more robust language model: a question-and-answer mechanism plus an attentive layer responsible for handling the increased contextual demands of their corpus (Wilmot and Keller, 2021). They contribute an original learning pipeline that can identify salient sentences across a corpus of fiction. Additionally, Ted Underwood's work studying how to measure narrative time pairs excellently with Genette's narratology (Underwood, 2018). For my interests, capturing the "amount" of time elapsed in a given passage would be a valuable data point for an such an analysis. The methods to produce Underwood's duration values rely on a refined annotation scheme, but there are syntactic and morphological markers of time passing (e.g. "the next day") to capture.

3 Data

3.1 Corpus Formulation, Passage Sampling, Annotation

I source novels from *Project Gutenberg* using the `c-w gutenberg` wrapper with a script that automatically grabs digital texts (Wolff, 2021; Gutenberg).[1] I am currently working with 28 novels. I wrote a sampling script that generates a random number of samples of a particular length from a corpus. Along these lines, I made a determination that 800 characters was the optimal length for each sample. I adopted this cutoff after a process of exploring passages of varying lengths.

From there, I use the sampling code to randomly draw a selection of 364 passages from my corpus. Next, I wrote a script that randomly pulls a sample from a directory and prompts the user to apply a rating. To affix the scores, I input a detail "rating" (scaled from 1.0 to 5.0) for the passage. The name of the work and author is not visible during this evaluation. As I've read almost all the works in my corpus (some more than once), I'm mostly aware of what book or author the sample originated from — especially when an icon like Mr. Darcy graces me with his presence. While considering a text's score, I paid close attention to the amount of distinct objects and places that were mentioned and described. Passages that clumsily summarize an event or consist of choppy, back-and-forth dialogue receive lower scores than a Linnaean description of a sea-creature courtesy of Jules Verne's engrossing ouevre. Meanwhile, a fine-grained characterization of a lampshade receives a high score. Interestingly, 4 of the randoms samples explicitly use the word {detail}. The authors of these passages seemingly hold a conversation with narratologists like Barthes and Genette. For example, a monologue in Fyodor Dostoyevsky *Crime and Punishment* muses about "the faces of clerks absorbed in petty details" (Dostoyevsky). Indeed, astute authors are quite aware of the tension between unravelling a plot and properly enmeshing a reader into the character's milieu. Bolstering a scene with rich, vivid imagery enhances the texture of an author's set-pieces and envelops a narrative core in a drape of fulsome, descriptive imagery.

4 Methods and Analysis

4.1 Exploring the Ratings

After reading and tagging each passage, I now have a (first-pass) rating for the level of "detail" they contain. The distribution of the scores is roughly binomial, which follows a general sensibility: passages are either pretty detailed, or not that detailed. The median rating is 3.3 and the mean rating is 3.1, indicative of a left-skew (as in, there are slightly more resoundingly "not" detailed passages compared to surefire detailed passages). There are 197 passages with ratings greater than the mean and 167 with ratings lower than the mean. While the ratings I generated are by no means perfect, they constitute a constructive first attempt to make sense of this group of passages. Based on each sample's relation to the mean (above or below), I thus construe a partition of pseudo "labels": `detail` and `not_detail`. These labels are used in later sections as a comparable binary during aggregations and graphing. Again, they are certainly not authoritative, but they represent the theoretical end-product

[1] The code mentioned throughout this project, along with program descriptions, an itemized corpus, the full feature-matrix, etc. can be viewed at github.com/timschott/details.

of a study premised on ratings sourced from repeatable, objective heuristics. The divisions are an imperfect yet positive first step towards a formulation of how to identify a fictional passage's details.

4.2 Log-Odds

The first coordinated experiment I run with my random samples implements the log-odds ratio with an informative prior (Monroe et al., 2008). The goal of this routine is to determine which words best represent my two pseudo-camps — as used to aplomb in existing research(Monroe et al., 2008; Jurafsky et al., 2014). For every word in the corpus, I carry out a frequency-based analysis (both per group and overall) and output a score for each word that represents its allegiance with one of the two camps. As a prior probability, I use the word's overall frequency across the entire (unsampled) corpus. For example, the 3 most distinctive words aligned with `detail` samples are {the, of, with} while the 3 most aligned with `not_detail` samples are {you, is, be}. Interesting differences manifest from this cross-section of the corpus; the presence of the second-person "you" suggests an outsized frequency of exchanges between particular characters and therefore signals that the sample at hand contains sparse detail. To continue, the top-25 for the `detail` group contains words for colors {black, white}, numbers {two, three}, prepositions {through, behind}, naturalistic elements {wind, air} and man-made settings {house, room}. Different categories of details (colors, numbers, naturalistic phenomena) organically emerge as we contemplate this output. For instance, associating prepositions with detail passages comports with the preposition's central function of revealing the relationship between relationally-paired nouns.

4.3 Specificity, Parts-of-Speech

My next experiment quantifies the "specificity" found in each sample. Naturally, we could surmise that passages trafficking in details would use highly specific words. The question, of course, is how to systematically identify specific language. Nelson (2020) provides a method for operationalizing and measuring specificity. I use the `SpaCy` software package to part-of-speech tag and lemmatize an inputted sample (Honnibal and Montani, 2017). Then, for each (adjudged) noun and verb used in the sample, the distance from the word's lemmatized form to its broadest hypernym is calculated with the `nltk` Wordnet API (Toolkit, 2021; University). Per Nelson (2020), I then average these distance calculations for each random sample. To simplify this procedure, I use the first (typically the most-frequent) synset found in WordNet (e.g. for house, `house.n.01`). This method is successful at roughly an equal rate to much more sophisticated word sense disambiguation schemes (Raganato et al., 2017). In this routine, I also record part-of-speech tallies for classes of interest: noun, verb, adjective, adverb and adpositions (a superset of prepositions).

In step with the log-odds calculations, inspections of the specificity results are quite enlightening. All 10 of the most specific samples are members of the `detail` group; 9 out of the 10 least specific samples belong to the `not_detail` group. The rogue sample originates from Daniel Defoe's *Robinson Crusoe*, whose earnest, idiosyncratic prose — its noun usage, especially — presents one of the more challenging strains of writing to classify (Defoe)[2]. The source of the most specific passage in the corpus is Tolstoy's *Anna Karenina*, when the author lyrically describes a passing crowd: "...all were flooded with light. on the right side of the warm church, in the crowd of frock coats and white ties, uniforms and broadcloth, velvet, satin, hair and flowers, bare shoulders..." (Tolstoy). A passage like this, equipped with conminuted nouns like {broadcloth} illustrates that the specificity metric is a helpful data point to consider when hunting for details. While readers obviously cannot recreate the query-based work used to produce this metric, they should be advised to heed the granularity of an author's nouns and verbs e.g. whether a character errands to a {store} or the more specific {fishmonger}.

To continue, `detail` samples are 14% more specific than `not_detail` samples. The part-of-speech can statistics explain this result. Overall, hypernym chain for any given verb is likely to be much smaller than that of a noun. A commonly used verb such as {enter} is the broadest member of its synset, while a commonly used noun like {house} is 9 levels down from its broadest member, {entity}. As such, a passage with more nouns than verbs will likely boast a high specificity score. On average,

[2]This brings to mind Virginia Woolf's memorable observation regarding the treatment of objects in the work: "Thus Defoe, by reiterating that nothing but a plain earthenware pot stands in the foreground, persuades us to see remote islands and the solitudes of the human soul."(Liu, 1999)

detail samples contain 17% more nouns than their counterparts and, likewise, 17% more adpositions. With this in mind, we can orient an analysis of the divergent subject matter and diction of these groups about a novel axis: the detail samples deliver information about things, places, and states of being while the not_detail passages narrate events and kinetic activities. Moving forward, I am excited to pursue related syntactic lines of inquiry. For instance, investigations into preposition usage possess potential for fertile results.

4.4 Measuring Time (Is Hard To Do)

One point of difference I speculate between detail and not_detail samples involves their conception of time. Recall how narratologists like Genette remark that authors can, as it were, "zoom-in" on a narrative set piece, thereby "suspending" the course of narrative time (Genette, 1976). Traditionally, one method of understanding the subject matter discussed across a corpus is to leverage an outside knowledge base or dictionary. In the digital humanities, the *Harvard Inquirer* has long-served as one such resource (Harvard, 2012). For example, a study seeking to reveal what makes a poem beautiful tallies the usage of words that are members of categories ranging from "concrete" to "psychological" (Kao and Jurafsky, 2012). Similar to my earlier success using WordNet, for each annotated sample I tabulated the frequency of words that fall into the *Inquirer*'s "time" category, which includes entries such as {before} and {early}.

Unlike the specificity work, there is not much of a difference between the two pseudo-categories. detail on average contain 6.8 time words per sample, while not_detail passages contain 6.54 time words per sample. Additionally, when it comes to extremes, the 10 samples that use the greatest and smallest amount of "time" words are mostly balanced in terms of their pseudo-labels. Furthermore, the distributions of "time" word usage across the binary do not elucidate anything particularly noteworthy. The standard deviations, for example, are almost identical. Frequency-based methods are clearly not sufficient to capture the complicated concept of "time" passing within a text sample. This particular task is likely exacerbated by my use of passages rather than full chapters; certain passages, when rent from their original context, don't possess any indicators for how much time is passing. Much like the literal usage of {detail} within my dataset, authors consciously attend to the circadian rhythms of their texts. However, they don't always make it explicit, especially in the modernist era that followed Marcel Proust's famous snack. (Doubrovsky and Bové, 1975). Indeed, it's hard to imagine an elegant computational solution to this aspect of the duration problem since the exclusion of a single word from a sample can completely transform the amount of "time" contained in a passage. All in all, the methodology (relying on human annotators) utilized by Underwood is much more sound than my salvo. However, looking forward to my next method of analysis, there is still much to be gleaned from frequency-based experiments.

4.5 Prepositions and Chi-Squared

I proceeded to explore the proportional usage of prepositions. To start, I calculated the observed frequencies of each adposition (excluding punctuation from the denominator) across my detail and not_detail samples. I discarded entries in my preposition counters that appeared less than five times as these malformed observations were symptomatic of SpaCy's tagging system misfiring. From there, I constructed a 2x2 matrix per adposition which tracked the amount of times the word does and does not appear in the two categories. The goal of this exercise was to determine, for each preposition, whether its proportional usage was independent of its allegiance within the "detail" binary.

The Chi-Squared test for independence was used to evaluate this hypothesis.[3] For words with expected frequencies of less than 5, Fisher's exact test was employed (Coult, 1965). The null hypothesis was rejected at the 0.05 significance level for just over half of observed prepositions. The words with statistically significant p-values include prepositions for establishing directional and locational relationships such as {down, in, through}. Moreover 6 of the prepositions such as {across, behind, beside} appear numerous times, *exclusively*, in the detail samples. These fascinating takeaways notwithstanding, it is important to consider one of this test's complicating factors: Chi-squared's assumption of word independence. This assumption can hamper the test's efficacy when investigating natural language due to the organic repetition of relevant activities or objects within a passage (Bam-

[3] Full Chi-Squared test output with test statistics, p-values, etc. can be viewed in the project repository

man, 2021a). Prepositions, however, serve as ubiquitous facets of language employed across genre and subject matter and therefore are more compatible with this assumption than other linguistic classes.

For example, the two samples featuring the largest number of adpositions, 28, are located in Gustave Flaubert's *Madame Bovary* and Defoe's *Robinson Crusoe* respectively. Two novels more divergent would be challenging to conjure, yet, Defoe and Flaubert cohere their fictional worlds through the consistent application of prepositions such {on, through, within}.[4] Consider this dazzling snippet from Flaubert, boasting a healthy 7 prepositions across its 33 words: "...the stove-pipe, in the shape of a palm-tree, spread its gilt leaves over the white ceiling, and near them, outside the window, in the bright sunshine, a little fountain gurgled in a white basin..." (Flaubert). Flaubert's prepositions orient readers and reify a fictional setting – precisely the mimetic labor discussed by scholars of realism (Auerbach, 2003).

This experiment does not give me license to summarily conclude that more prepositions results in a more detailed passage. However, I am able to conclude that the usage of prepositions warrants significant attention when deciding if a passage merits the "detailed" accolade. Without a doubt, there is a definitive connection between "verisimilitude" (and other nebulous concepts used uncritically when praising prose) and prepositions. You'd be hard pressed to emulate Flaubert without using {around, across, between}.

4.6 Word Embeddings and BERT Intepretation

The final method of analysis I carried out relies on the word embeddings and large language models. This work enhanced my study with a nuanced, state-of-the-art strategy towards identifying details in samples from literature. Indeed, I wanted to match the successful efforts completed by computational narratologists. As previously mentioned, recent work in this domain includes fine-tuning a language model and running a clustering algorithm because salient sentences ought to cluster closely to an *a priori* salient summarization (Wilmot and Keller, 2021). In my case, I am interested in two aspects of language modeling: word embeddings and attention.

Word embeddings aim to transform words into numerical representations which captures syntactic and semantic meaning (Lin et al., 2016). These representations generate a column vector for every word across a corpus vocabulary. The dimensions of the matrix depend on your data. The row count equals the number of tokens in the corpus. The column count stores how large the embeddings are for a single token (i.e. "dimensionality") (Jurafsky and Martin, 2021). Every embedding implementation relies on the "distributional" hypothesis: similar words appear in similar contexts (Mikolov et al., 2013; Peters et al., 2018). The embeddings generated as part of the **BERT** model's initial training phase have offered a fruitful point of analysis in areas such as machine translation (Zhu et al., 2020). To this end, I set out to experiment in my problem space. My goal was not to produce yet another permutation of **BERT**; we have enough acronyms to keep track of. Rather, I'm interested in interpreting a subset of the model's output: its attention weights.

To start, I ran a simple classification trial. I trained off-the-shelf **BERT** models from the `transformers` library (Wolf et al., 2020). The best performing iteration was **BERT Medium** (named, of course, for its parsimonious 41.7 *million* parameters) which achieved a development-set accuracy of 80% on my annotated data (Turc et al., 2019). Considering I did not augment the base model or manipulate my dataset, this performance is quite impressive.[5]. From here, I dove deeply into one of **BERT**'s central learning mechanisms: attention (Vaswani et al., 2017; Galassi et al., 2021). Attention, in short, provides a greater number of parameters to optimize during the language model's process of learning what words ought to be arranged near one another and the "meaning" behind their representations (Bamman, 2021b). Across each sample, the model incorporates the sum of each token's attention scores into its classification decision. By inspecting the attention "weights" connected to each token following a classification exercise, we can approximate an understanding of how consequential a given token was when the model computed a decision boundary between the two classes.

I used the `PyTorch captum` library to explore

[4] In a nice gesture of consistency, both samples align with the `detail` camp.

[5] I ran a similar classification exercise using k-means and barely cracked 50% accuracy. Coin-flip territory.

the attention mechanism of my trained **BERT** classifier (Paszke et al., 2019). I aggregated each token's attention weight across every sample in the development dataset. In this exercise, a token with a positive aggregate aligns closer to `detail` samples. Next, I produced a look-up table with an entry for each of the 2770 distinct tokens found in the corpus. Naturally, some slots of the top-100 most positively "attended-to" tokens are occupied by punctuation, suffix fragments, artifacts of **BERT**'s tokenization scheme, etc (Devlin et al., 2019). However, on closer inspection, there are 9 prepositions found in the top-100, Of these prepositions, the majority {across, against, at, behind, of} boast statistically significantly scores from the earlier $\tilde{\chi}^2$ work. This is a fascinating reinforcement of my earlier experiments revealing `detail` passages rely on prepositions. The bottom-100 list is characterized by its own cluster of (conventionally-labeled) "stopwords" {it, was, the} pronouns {he, she, him}, indicators of spoken word {said, communication, accents} and auxiliaries {did, had}. Incidentally, this grouping contains only 6 prepositions, half of which boast statistically significantly $\tilde{\chi}^2$ scores. Besides underscoring my earlier findings regarding prepositions, this experiment sheds light on the outsized "attention" the language model places on facets of language like punctuation marks that a human reader might otherwise overlook. This highlights the tension and potential pitfalls that stem from uncritically relying on **BERT**'s classification decisions — particularly when analyzing literature, where language itself is a permanent fixture of discourse, or attempting to transfer language models across cultures and time periods. (Black, 1979; Kusnír, 2010; Cowart, 2012; Ruder et al., 2019),

5 Future Work

Looking ahead, learning strategies, like topic modeling, could certainly be of service in the task of identifying details (Klein, 2020; Grimmer, 2010)

Classification exercises, more broadly, will be an interesting mode of analysis once more samples have been annotated. It is quite encouraging that a relatively small data set high performance during classification exercises with **BERT** and there is clearly exciting potential for continued application of large language models in this territory. Additionally, I still think there is more to uncover concerning duration and the elapsing of "time" in texts. I conjecture that some flavor of "transfer learning," applying discoveries from one domain to another, could be of service here (Ruder et al., 2019). Training models using texts from an area like screenwriting and porting them onto literary datasets could generate valuable knowledge about details and their usages in literary and non-literary contexts

To continue, I would like to leverage my bank of metrics as the basis for formulating detail identification guidelines. Privacy scholars provide an excellent model for this activity when they fashion detailed rubrics for interrogating the compliance of terms-of-service agreements with respect to extant privacy law — a semantically challenging task not unlike mine (of Education, 2015). My goal is to draft a document of similar legibility and lucidity with fictional details at the core. A dedicated foray into this line of work is out of the scope of this current project, so it will be pursued in a separate line of inquiry. These guidelines will yield numerous avenues for further study and help answer important questions regarding how effectively humans and our distributed technological systems can recognize, quantify, and understand the details used in a novel. Additionally, this would also produce a dataset that could be evaluated through the use of Cohen's Kappa or other statistic, of participants that read the same passage with my guidelines as a reference (McHugh, 2012). The criteria will of course be tweaked, iterated and remolded based on the helpful recommendations of other like-minded scholars interested in applying computational modes of study to literary phenomena.

6 Conclusion

My work produces actionable insights regarding the usage of details in fiction. I show the importance of prepositions and nouns when tasked with identifying details. I also demonstrate the advantages and disadvantages of frequency-based analysis. Lastly, I amplify the exciting work of digital narratologists and present a novel integration of large language models that corroborates this study's frequency-based results. 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, to create the conditions in which constructing such a model is for the first time a real possibility.

References

Erich Auerbach. 2003. *Mimesis: the representation of reality in Western literature*. Princeton University Press.

Elaine Auyoung. 2015. Rethinking the Reality Effect. ISBN: 9780199978069.

David Bamman. 2021a. Distinctive Words.

David Bamman. 2021b. Info 256 lecture 15: Attention/bert.

Roland Barthes. 1989. *The Rustle of Language*. University of California Press.

Roland Barthes and Lionel Duisit. 1975. An Introduction to the Structural Analysis of Narrative. *New Literary History*, 6(2):237–272. Publisher: Johns Hopkins University Press.

Max Black. 1979. Wittgenstein's Language-games. *Dialectica*, 33(3/4):337–353. Publisher: Wiley.

Allan D. Coult. 1965. A Note on Fisher's Exact Test. *American Anthropologist*, 67(6):1537–1541. Publisher: [American Anthropological Association, Wiley].

David Cowart. 2012. *Don DeLillo: The Physics of Language*. University of Georgia Press, Athens.

Daniel Defoe. *The Life and Adventures of Ronbinson Crusoe*. Seeley, Service & Co. Limited, London.

Jacob Devlin, Ming-Wei Chang, Kenton Lee, and Kristina Toutanova. 2019. BERT: Pre-training of deep bidirectional transformers for language understanding. In *Proceedings of the 2019 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies, Volume 1 (Long and Short Papers)*, pages 4171–4186, Minneapolis, Minnesota. Association for Computational Linguistics.

Fyodor Dostoyevsky. *Crime and Punishment*.

Serge Doubrovsky and Carol Bové. 1975. The Place of the Madeleine: Writing and Phantasy in Proust. *boundary 2*, 4(1):107–134. Publisher: Duke University Press.

U.S. Department of Education. 2015. Protecting Student Privacy While Using Online Educational Services: Model Terms of Service. *U.S. Department of Education*, page 9.

Gustave Flaubert. *Madame Bovary*.

Andrea Galassi, Marco Lippi, and Paolo Torroni. 2021. Attention in Natural Language Processing. *IEEE Transactions on Neural Networks and Learning Systems*, 32(10):4291–4308. ArXiv: 1902.02181.

Gérard Genette. 1976. Boundaries of Narrative. *New Literary History*, 8(1):1–13. Publisher: Johns Hopkins University Press.

Justin Grimmer. 2010. A Bayesian Hierarchical Topic Model for Political Texts: Measuring Expressed Agendas in Senate Press Releases. *Political Analysis*, 18(1):1–35.

Project Gutenberg. Project Gutenberg.

Harvard. 2012. Harvard Inquirer.

Matthew Honnibal and Ines Montani. 2017. spaCy 2: Natural language understanding with Bloom embeddings, convolutional neural networks and incremental parsing. To appear.

Dan Jurafsky, Victor Chahuneau, Bryan R. Routledge, and Noah A. Smith. 2014. Narrative framing of consumer sentiment in online restaurant reviews. *First Monday*.

Dan Jurafsky and James Martin. 2021. Vector Semantics and Embeddings. In *Speech and Language Processing*.

Justine Kao and Dan Jurafsky. 2012. A Computational Analysis of Style, Affect, and Imagery in Contemporary Poetry. page 10.

Judith Klein. 2020. Use MT to simplify and speed up your alignment for TM creation. In *Proceedings of the 14th Conference of the Association for Machine Translation in the Americas (Volume 2: User Track)*, pages 237–269, Virtual. Association for Machine Translation in the Americas.

Jaroslav Kusnír. 2010. Stephen J. Burn, Jonathan Franzen at the End of Postmodernism. *European journal of American studies*. Publisher: European Association for American Studies.

Yankai Lin, Shiqi Shen, Zhiyuan Liu, Huanbo Luan, and Maosong Sun. 2016. Neural Relation Extraction with Selective Attention over Instances. In *Proceedings of the 54th Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers)*, pages 2124–2133, Berlin, Germany. Association for Computational Linguistics.

Lydia H. Liu. 1999. Robinson Crusoe's Earthenware Pot. *Critical Inquiry*, 25(4):728–757. Publisher: The University of Chicago Press.

Genevieve Liveley. 2019. Russian formalism. In *Narratology*. Oxford University Press, Oxford.

Mary L. McHugh. 2012. Interrater reliability: the kappa statistic. *Biochemia Medica*, 22(3):276–282.

Tomas Mikolov, Ilya Sutskever, Kai Chen, Greg S Corrado, and Jeff Dean. 2013. Distributed Representations of Words and Phrases and their Compositionality. In *Advances in Neural Information Processing Systems*, volume 26. Curran Associates, Inc.

David H. Miles. 1979. Reality and the Two Realisms: Mimesis in Auerbach, Lukács, and Handke. *Monatshefte*, 71(4):371–378. Publisher: University of Wisconsin Press.

Burt L. Monroe, Michael P. Colaresi, and Kevin M. Quinn. 2008. Fightin' Words: Lexical Feature Selection and Evaluation for Identifying the Content of Political Conflict. *Political Analysis*, 16(4):372–403.

Laura K. Nelson. 2020. Computational Grounded Theory: A Methodological Framework. *Sociological Methods & Research*, 49(1):3–42. Publisher: SAGE Publications Inc.

Takaki Otake, Sho Yokoi, Naoya Inoue, Ryo Takahashi, Tatsuki Kuribayashi, and Kentaro Inui. 2020. Modeling event salience in narratives via barthes' cardinal functions. In *Proceedings of the 28th International Conference on Computational Linguistics*, pages 1784–1794, Barcelona, Spain (Online). International Committee on Computational Linguistics.

Adam Paszke, Sam Gross, Francisco Massa, Adam Lerer, James Bradbury, Gregory Chanan, Trevor Killeen, Zeming Lin, Natalia Gimelshein, Luca Antiga, Alban Desmaison, Andreas Kopf, Edward Yang, Zachary DeVito, Martin Raison, Alykhan Tejani, Sasank Chilamkurthy, Benoit Steiner, Lu Fang, Junjie Bai, and Soumith Chintala. 2019. Pytorch: An imperative style, high-performance deep learning library.

Matthew E. Peters, Mark Neumann, Mohit Iyyer, Matt Gardner, Christopher Clark, Kenton Lee, and Luke Zettlemoyer. 2018. Deep contextualized word representations. *arXiv:1802.05365 [cs]*. ArXiv: 1802.05365.

Alessandro Raganato, Claudio Delli Bovi, and Roberto Navigli. 2017. Neural Sequence Learning Models for Word Sense Disambiguation. In *Proceedings of the 2017 Conference on Empirical Methods in Natural Language Processing*, pages 1156–1167, Copenhagen, Denmark. Association for Computational Linguistics.

Sebastian Ruder, Matthew E. Peters, Swabha Swayamdipta, and Thomas Wolf. 2019. Transfer Learning in Natural Language Processing. In *Proceedings of the 2019 Conference of the North American Chapter of the Association for Computational Linguistics: Tutorials*, pages 15–18, Minneapolis, Minnesota. Association for Computational Linguistics.

Naomi Schor. 1984. Details and Realism: Le Curé de Tours. *Poetics Today*, 5(4):701–709. Publisher: [Duke University Press, Porter Institute for Poetics and Semiotics].

Leo Tolstoy. *Anna Karenina*.

Natural Language Toolkit. 2021. Nltk. Original-date: 2009-09-07T10:53:58Z.

Iulia Turc, Ming-Wei Chang, Kenton Lee, and Kristina Toutanova. 2019. Well-Read Students Learn Better: On the Importance of Pre-training Compact Models. *arXiv:1908.08962 [cs]*. ArXiv: 1908.08962.

Ted Underwood. 2018. Why Literary Time is Measured in Minutes. *ELH*, 85(2):341–365.

Princeton University. Princeton University "About WordNet.".

Ashish Vaswani, Noam Shazeer, Niki Parmar, Jakob Uszkoreit, Llion Jones, Aidan N. Gomez, Lukasz Kaiser, and Illia Polosukhin. 2017. Attention Is All You Need. *arXiv:1706.03762 [cs]*. ArXiv: 1706.03762.

David Wilmot and Frank Keller. 2020. Modelling Suspense in Short Stories as Uncertainty Reduction over Neural Representation. *arXiv:2004.14905 [cs]*. ArXiv: 2004.14905.

David Wilmot and Frank Keller. 2021. Memory and Knowledge Augmented Language Models for Inferring Salience in Long-Form Stories. *arXiv:2109.03754 [cs]*. ArXiv: 2109.03754.

Thomas Wolf, Lysandre Debut, Victor Sanh, Julien Chaumond, Clement Delangue, Anthony Moi, Perric Cistac, Clara Ma, Yacine Jernite, Julien Plu, Canwen Xu, Teven Le Scao, Sylvain Gugger, Mariama Drame, Quentin Lhoest, and Alexander M. Rush. 2020. Transformers: State-of-the-Art Natural Language Processing. Pages: 38–45 original-date: 2018-10-29T13:56:00Z.

Clemens Wolff. 2021. c-w-gutenberg. Original-date: 2021-03-29T04:32:36Z.

Jinhua Zhu, Yingce Xia, Lijun Wu, Di He, Tao Qin, Wengang Zhou, Houqiang Li, and Tie-Yan Liu. 2020. Incorporating BERT into Neural Machine Translation. *arXiv:2002.06823 [cs]*. ArXiv: 2002.06823.