

Adjective Analysis Between Two Writers of the Victorian Era

Alessandro Gubler S5994861

Length: 3 pages

Abstract

The goal of this report is to establish the key differences in character descriptions between two writers within the Victorian era, namely Mark Twain and Charles Dickens. We will analyze and compare texts from both authors found in Project Gutenberg, which is a digital archive of culturally important books. In this research, we will be comparing the frequency of adjectives and the emotional tone used in character descriptions to determine the difference between these two writers. We aim to uncover how their use of language impacts their narrative style of writing. This research contributes to the study of language variation by showing how important adjectives and emotional tone are in shaping characters.

1 Introduction

Language is the most powerful tool for making people feel emotions and creating iconic characters in literature that people resonate with. With that in mind, the stylistic choices authors make, such as the use of adjectives, are very important in shaping how a character develops over the course of a book. This study explores the difference between two iconic authors, Charles Dickens and Mark Twain, namely in how they use adjectives to describe characters. We believe this research will help us better understand language variation in the Victorian era of literature. Charles Dickens is known for his elaborate storytelling and often detailed portrayals of characters, using intricate language to create iconic characters. Mark Twain, however, is known for his usage of humor in language and his social critique, which means that his language has a more direct and conversa-

tional style, leading to perhaps less iconic characters. Henceforth, comparing their use of adjectives will provide us with a proper insight into how these two writers compare against each other. Twain offers a completely unique writing style in his time, as McKay (1976) observes in her analysis of the famous novel *The Adventures of Huckleberry Finn*. Twain uses an ordinary, conversational style of language while incorporating humor and social critique, which is rather innovative and unique for writers in this era. According to Gardner (1969), something of importance to note is that Twain did not appreciate the writings of Dickens and actively avoided his influence. There is actually no evidence that either writers were influenced by one another, even though they were often exposed to each other's writing. Even though there's plenty of research to be found on Charles Dickens and Mark Twain, there's nowhere near enough research on their contrasting narrative styles, the role of adjectives in character descriptions, and how these impact storytelling. In this research paper, we will answer the following question: How does the use of adjectives compare between Mark Twain and Charles Dickens in their descriptions of characters? We hypothesize that Charles Dickens will use more adjectives per 100 words in character descriptions than Mark Twain does, and the tone of these adjectives will be more nuanced or dramatical, while Mark Twain will use more neutral and humorous adjectives.

2 Related Work

While there are a decent amount of studies and articles on everything concerning Mark Twain and Charles Dickens, as well as other relevant parts of this study such as sentiment analysis, we don't believe adjective frequency relating to character descriptions is something that's been done often be-

fore. There is a very limited amount of research to be found on the use of adjectives in character descriptions. A study we found very interesting and most relevant is Gardner (1969), because it provides a detailed insight in the background between Mark Twain and Charles Dickens, which can help us figure out what makes it so there's an actual difference between the two writers. Furthermore, McKay (1976) analyses a famous work by Mark Twain, namely *Adventures of Huckleberry Finn*, which gives us an overview of Twain's unique and innovative writing style. Lastly, we make use of the work of Samir et al. (2021) who have done elaborate research on sentiment analysis using Twitter tweets and the language model BERT, which is what we will be using to do our sentiment analysis with as well. This article provides us with a very useful insight on how sentiment analysis actually works, and how we can use BERT to do our sentiment analysis.

3 Data

For this research, we will be using texts written by Charles Dickens and Mark Twain, found in the public digital archive Project Gutenberg. We will be using novels with as many character descriptions as possible, namely Dickens' *Great Expectations* and *Oliver Twist*, and Twain's *The Adventures of Huckleberry Finn* and *The Adventures of Tom Sawyer*. These novels should give us plenty of material to analyze character descriptions with based on adjective use. Our independent variable is the author of the book. This will either be Charles Dickens or Mark Twain. Our dependent variables are about the use of adjectives. For adjective frequency: we will extract adjectives from character descriptions using a natural language processing (NLP) pipeline like spaCy. For measuring the adjective tone, we will utilize a sentiment analysis tool from Hugging Face, inspired by Samir et al. (2021) to determine the emotional tone (positive, neutral, or negative) of adjectives. This model is a modified version of BERT and is trained on 124 million tweets ranging from January 2018 to December 2021. An overview can also be found on our GitHub.

Pre-processing Firstly, we will grab our required novels from Project Gutenberg as a PDF. We will then tokenize all our data by splitting sentences into words using the spaCy NLP library. We will be removing non-informative words like

“the” and “and”, in order to maximize our adjective efficiency. To actually isolate the adjectives and get the information we want, we will be using part-of-speech tagging together with character identification to filter out sentences that describe characters we have collected information on beforehand. Sentences that are not relevant to character descriptions and do not add anything of value will be excluded with automated filters. The isolated adjectives will then be fed into the sentiment analysis tool and from that we have a good starting point. There are of course some limitations to this approach, for example bias in text selection. By only choosing to analyse a few novels by an author, we might not actually grasp the full range of the author's writing style and will thus be unrepresentative. We also have to be careful with using the sentiment analysis tool, as it is trained on more recent data. It may misinterpret the actual tone of adjectives because it is missing context and the text often contains more dated language.

Table 1 provides an example table of the data that will be used in this study.

	Twain Novels	Dickens Novels
Adjective frequency	x	x
Adjective tone	x	x

Table 1: Overview of the data set. After collecting results, the x's will be replaced with actual values.

4 Predicted Results

Based on our hypothesis, we do expect to find some differences in the use of adjectives between Charles Dickens and Mark Twain, for example in the adjective tone. We expect that adjectives in Dickens' novels will be more pronounced with the goal of conveying heavy emotions. This fits in place with his more theatrical and elaborate narrative style. Conversely, adjectives in Twain's novels are going to be more neutral, reflecting his conversational and storytelling. We also expect to find a difference in adjective frequency. We expect that Dickens' texts contain a higher frequency of adjectives, because of the detailed character portrayals we spoke about before. Twain's texts, however, are expected to have fewer adjectives overall, focusing instead on straightforward and simple descriptions everyone could understand.

Table 2 provides an example table of the result after analysing two novels.

Adjective	Twain Novels	Dickens Novels
Frequency	3.2	5.8
Tone	Neutral	Positive

Table 2: Overview of the data set. The values are purely made up and should not be taken into any calculations

Discussion If our expected results are to be true, this study will have provided insight into the difference between the linguistic strategies employed by Mark Twain and Charles Dickens. This study will provide further insight into how useful the adjective is for understanding linguistic descriptions of characters. These findings will also show how social aspects impact the view of an author on storytelling methods. However, it should be noted that this study is extremely small and has a very little sample size. If you wanted to properly analyse the tone of character descriptions, we would recommend looking at more linguistic devices than only the adjective, as it is not representative enough. We would also want to look more at the social and cultural elements that take place in the background of the author’s life, to understand how the author has developed their worldview and thus narrative style.

5 Conclusion

This study aimed at exploring the language variation between two Victorian era writers, Mark Twain and Charles Dickens, shows that there is a clear difference in their use of adjectives when it comes to describing characters. This research revealed that Dickens has a more frequent and detailed use of adjectives, and Twain has a more neutral, straightforward and humorous approach to using adjectives to describe characters. Our research is limited by the small sample size and lack of context by only looking at the adjective, we might not be able to fully represent the style of both authors with only two novels. The reliance on automated tools may also prove inaccurate, considering they were trained on more recent data and the novels might use archaic language. Overall, this study contributes to the field of linguistics by illustrating how adjective use can shape the portrayal of characters and impact storytelling.

References

- Gardner, J. H. (1969). Mark twain and dickens. *PMLA* 84(1), 90–101.
- McKay, J. H. (1976). "tears and flapdoodle": Point of view and style in "the adventures of huckleberry finn". *Style* 10(1), 41–50.
- Samir, A., S. M. Elkaffas, and M. M. Madbouly (2021). Twitter sentiment analysis using bert. In *2021 31st International Conference on Computer Theory and Applications (ICCTA)*, pp. 182–186.