Text Analysis Report: modernist_literature_chatgpt

1. Basic Statistics

Metric	Value
Word Count	753
Sentence Count	36
Character Count	5520
Unique Word Count	170
Avg Word Length (chars)	6.13
Avg Sentence Length (words)	20.92
Avg Sentence Length (chars)	153.33
Sentence Length Std Dev (words)	9.91
Lexical Diversity (TTR)	0.2258
MATTR (Window=50)	0.8624

2. Entropy & Complexity
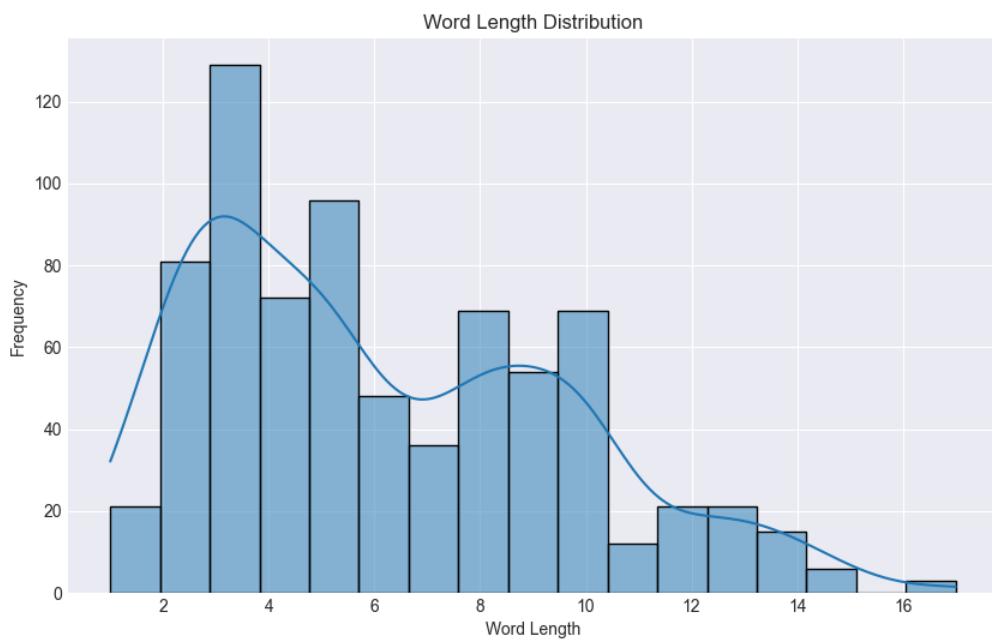
Metric	Value
Shannon Entropy (Words)	6.9751 bits
Shannon Entropy (Chars)	4.5925 bits
Function Word Ratio	0.2948

3. Readability Scores
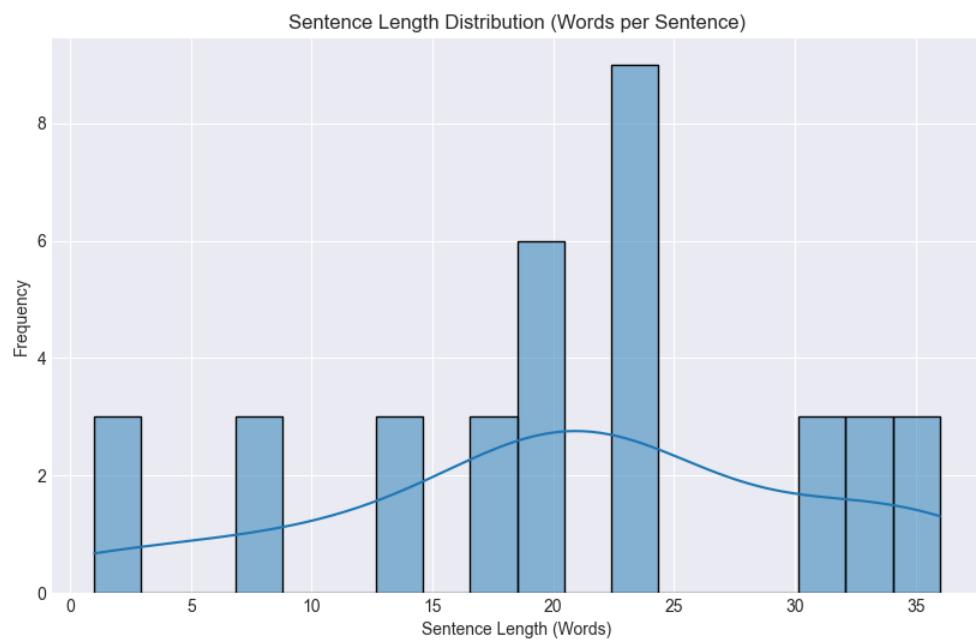
Index	Score
Flesch Reading Ease	6.64
Flesch-Kincaid Grade Level	17.80
Gunning Fog Index	12.10
SMOG Index	18.10
Coleman-Liau Index	19.26
Automated Readability Index	20.40

4. Visualizations

Word Length Distribution

Word Length Distribution

Sentence Length Distribution
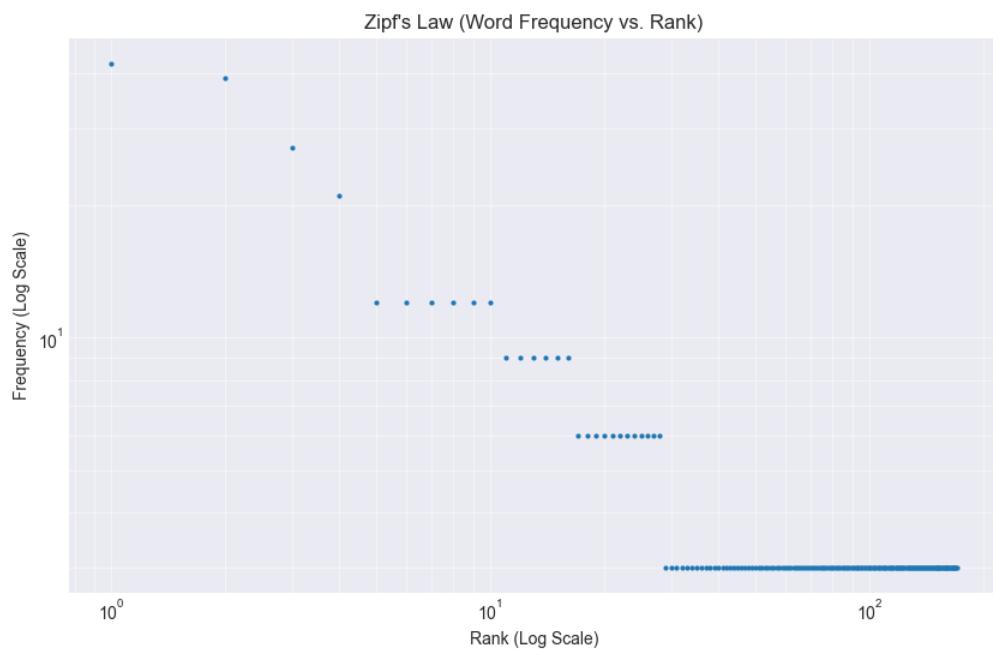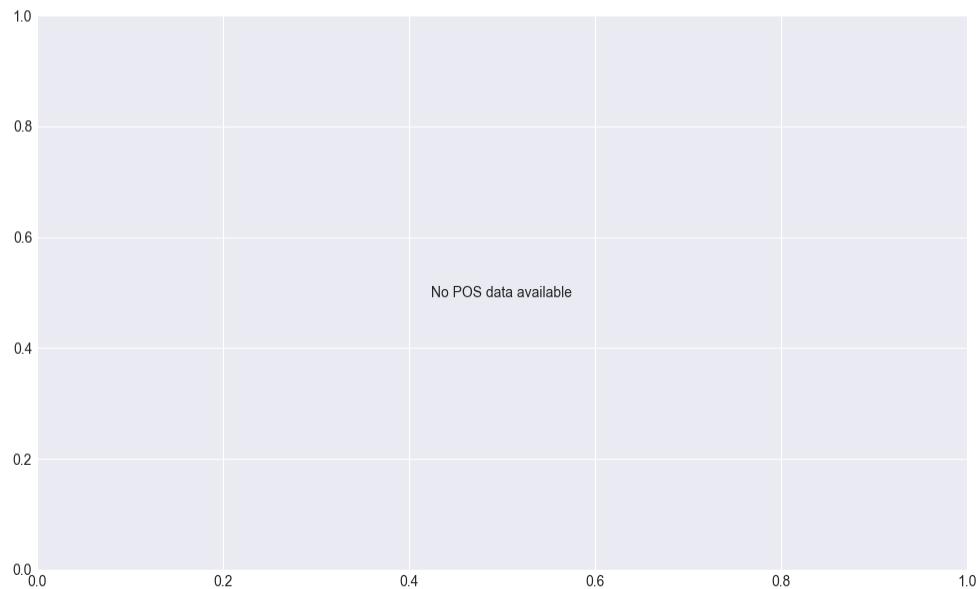
Sentence Length Distribution (Words per Sentence)
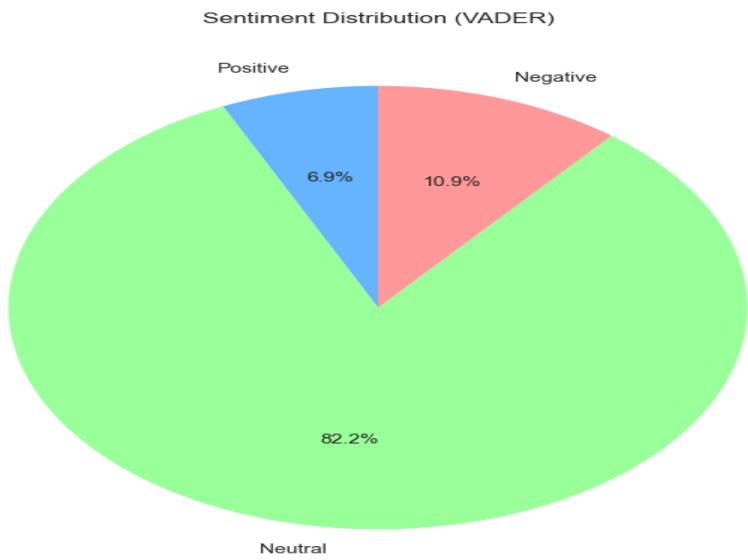
Zipf's Law Analysis (Word Frequency vs Rank)

Zipf's Law (Word Frequency vs. Rank)
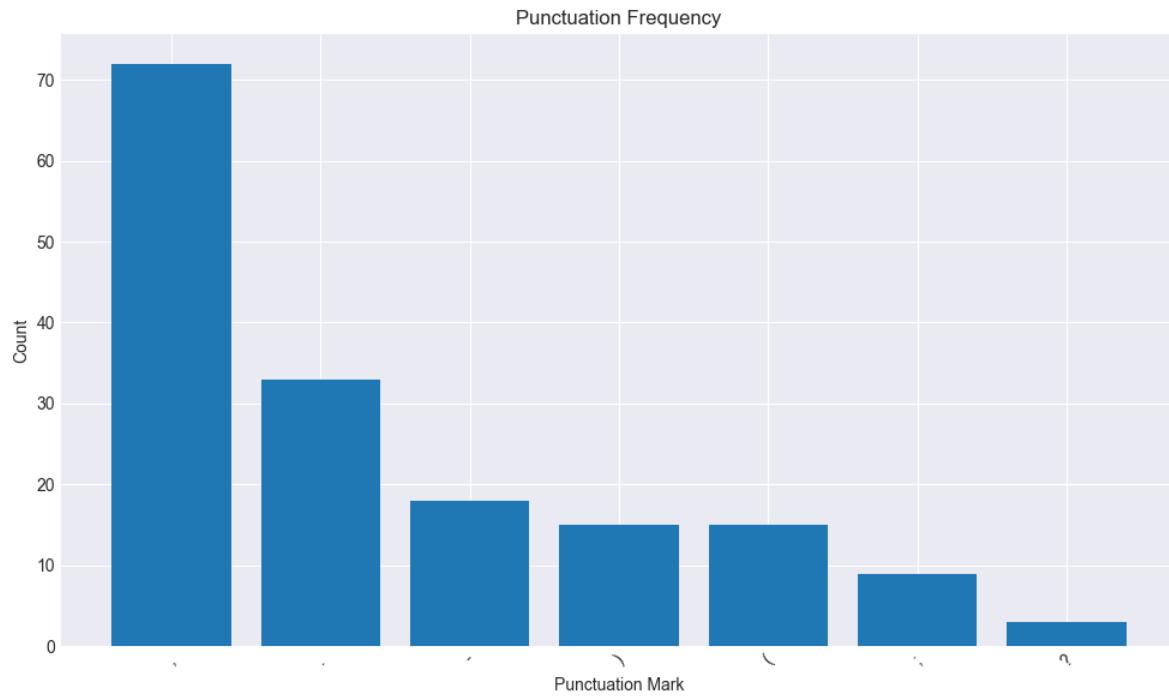
## Part-of-Speech Distribution (Top 25)

No POS data available

Sentiment Distribution (VADER)

Sentiment Distribution (VADER)

Punctuation Frequency

5. Frequency Analysis

Most Common Words (Top 15, with stopwords):

and (42), the (39), of (27), in (21), literary (12), modernism (12), a (12), to (12), consciousness (12), literature (12), movement (9), by (9), world (9), war (9), fragmented (9)

Most Common Punctuation:

',' (72), '.' (33), '-' (18), '(' (15), ')' (15), ';' (9), '?' (3)

6. N-gram Analysis

Common 2-grams (Top 10):

'literary modernism' (9), 'and the' (9), 'stream of' (9), 'of consciousness' (9), 'and a' (6), 'world war' (6), 'consciousness techniques' (6), 'the movement' (6), 'modernism literary' (3), 'modernism was' (3)
2-gram Repetition Rate: 0.6822

Common 3-grams (Top 10):

'stream of consciousness' (9), 'of consciousness techniques' (6), 'literary modernism literary' (3), 'modernism literary modernism' (3), 'literary modernism was' (3), 'modernism was a' (3), 'was a movement' (3), 'a movement that' (3), 'movement that emerged' (3), 'that emerged in' (3)
3-gram Repetition Rate: 0.6698

Common 4-grams (Top 10):

'stream of consciousness techniques' (6), 'literary modernism literary modernism' (3), 'modernism literary modernism was' (3), 'literary modernism was a' (3), 'modernism was a movement' (3), 'was a movement that' (3), 'a movement that emerged' (3), 'movement that emerged in' (3), 'that emerged in the' (3), 'emerged in the late' (3)
4-gram Repetition Rate: 0.6667

Character N-gram Analysis:

Common Character 3-grams (Top 10):

```
'ent' (51), '_in' (51), 'nd_' (51), '_an' (48), '_th' (45), 'the' (45), 'and' (45), 'ed_' (42),
'ter' (39), 'he_' (39)
```

Common Character 4-grams (Top 10):

```
'and_' (45), '_and' (42), '_the' (39), 'the_' (39), 'ment' (33), 'mode' (30), 'oder' (30),
'dern' (30), 'tion' (30), 'iter' (27)
```

Common Character 5-grams (Top 10):

```
'_and_' (42), '_the_' (36), 'moder' (30), 'odern' (30), 'liter' (24), 'itera' (24), 'derni'
(24), 'ernis' (24), '_mode' (21), 'ation' (21)
```