Mining Rhetorical Devices by means of Natural Language Processing

Chair of Web Technology and Information Systems
Prof. Dr. Benno Stein

Advisor: Khalid Al-Khatib

Bauhaus-Universität Weimar

Master Thesis Defense
January 23rd, 2018
What is Rhetoric?
What is Rhetoric?
What is Rhetoric?

Bob
What is a Rhetorical Device?

...40% whiter than the other leading brands

This is a must-see movie If you buy it now...

...the biggest, the brightest and the best.
Classification

Rhetoric

- Invention
- Arrangement
- Style
- Memory
- Delivery
Classify:

Rhetoric

- Invention
- Arrangement
- Style
- Memory
- Delivery

Bob:

- Open happiness.
- Grab a bottle.
- Don’t be sad!

Happy:

- Happiness
- Sad
- Bottle

Feelings:

- Drink
Classification

Invention | Arrangement | Style | Memory | Delivery

Rhetoric

Bob

Don’t be sad! Grab a bottle. Open happiness.
Classification

Rhetoric

Invention
Arrangement
Style
Memory
Delivery

Don’t be sad! Grab a bottle. Open happiness.

Feeling down? Open a bottle, open happiness!

Rhetorical question
Repetition, Balance

Bob
Bob

Feeling down? Open a bottle, open happiness!
Feeling down? Open a bottle, open happiness!
Feeling down? Open a bottle, open happiness!
Classification

Rhetoric

- Invention
- Arrangement
- Style
- Memory
- Delivery

Feeling down? Open a bottle, open happiness!

Bob
Classification

Rhetoric

Invention  Arrangement  Style  Memory  Delivery
Classification

Rhetoric

- Invention
- Arrangement
- Style
- Memory
- Delivery

Rhetorical Devices

- Schemes (syntax)
- Tropes (semantic)

Figures of thought
Envisioned Applications

Rhetoric-based NLG system
Envisioned Applications

Rhetorical style suggestion system

- **text (input)**
- **detection**
- **identification of rhetorical devices**
- **analysis**
- **rhetorical patterns**
- **suggestion**
- **suggestion (output)**
Research Questions
Research Questions
Research Questions
Research Questions

Genres

Topics

Authors
Detection of Rhetorical Devices
Pipeline – UIMA
Pipeline – UIMA
Pipeline – UIMA
Pipeline – UIMA Ruta

- **UIMA Rule-based Text Annotation** - intuitive and flexible domain specific language for defining patterns of annotations (Klügl et al. [2016]).

- Example:

```
DECLARE Sentence;
PERIOD #{} MARK(Sentence) PERIOD;
```

...This is a sample sentence...
Pipeline – Stanford CoreNLP

• **Stanford CoreNLP** – a suite of tools for linguistic analysis.

• We use:
  - Stanford Parser
  
  ```
  My dog also likes eating sausage.
  ```
  
  - Stanford Dependencies
  
  ```
  My dog also likes eating sausage.
  ```
Pipeline – UIMA
Pipeline – UIMA
But I must explain to you how all this mistaken idea of denouncing pleasure and praising pain was born and I will give you a complete account of the system, and expound the actual teachings of the great explorer of the truth, the master-builder of human happiness. No one rejects, dislikes, or avoids pleasure itself, because it is pleasure.
Rhetorical Devices

- Balance schemes: Interplay between equivalent ideas
- Omission schemes: Control the rhythm of thought
- Repetition schemes
- Custom schemes
Rhetorical Devices

Balance schemes
- Interplay between equivalent ideas
- Control the rhythm of thought

Omission schemes
- Deliberate omission of intuitive words
- Cause incompleteness

Repetition schemes

Custom schemes
Rhetorical Devices

Balance schemes
- Interplay between equivalent ideas
- Control the rhythm of thought

Omission schemes
- Deliberate omission of intuitive words
- Cause incompleteness

Repetition schemes
- Repetition of key words/ideas
- Used for emphasis or amplification
- Key to persuasion (according to Aristotle)

Custom schemes
Rhetorical Devices

Balance schemes
- Interplay between equivalent ideas
- Control the rhythm of thought

Omission schemes
- Deliberate omission of intuitive words
- Cause incompleteness

Repetition schemes
- Repetition of key words/ideas
- Used for emphasis or amplification
- Key to persuasion (according to Aristotle)

Custom schemes
- Informal rhetorical devices
- Strong emotional effect
- Includes causality, comparatives and voice
### Rhetorical Devices

#### Balance schemes
- Enumeration
- Pysma
- Isocolon
  - bicolon
  - tricolon
  - tetracolon

#### Omission schemes
- Asyndeton
- Hypozeugma
- Epizeugma

#### Repetition schemes
- Epanalepsis
- Mesarchia
- Epiphoza
- Mesodiplosis
- Anadiplosis
- Diacope
- Epizeuxis
- Polysyndeton

#### Custom schemes
- If-conditional 0
- If-conditional 1
- If-conditional 2
- If-conditional 3
- If-counterfactual
- Unless-cond.
- Whether-cond.
- Comparative Adjectives/Adverbs
- Superlative Adjectives/Adverbs
Rhetorical Devices

Balance schemes
- Enumeration
- Pysma
- Isocolon
  - bicolon
  - tricolon
  - tetracolon

Omission schemes
- Asyndeton
- Hypozeugma
- Epizeugma

Repetition schemes
- Epanalepsis
- Mesarchia
- Epiphaza
- Mesodiplosis
- Anadiplosis
- Diacope
- Epizeuxis
- Polysyndeton

Custom schemes
- If-conditional 0
- If-conditional 1
- If-conditional 2
- If-conditional 3
- If-counterfactual
- Unless-cond.
- Whether-cond.
- Comparative Adjectives/Adverbs
- Superlative Adjectives/Adverbs
**Balance: Enumeration**

**Enumeration** - a rhetorical device used to list a series of details, words or phrases. (literarydevices.net)

Old farmer had a pig, a dog, a cow and a horse.

Old farmer had a pig, a dog, a cow and a horse.
Balance: Enumeration

**Enumeration** - a rhetorical device used to list a series of details, words or phrases. ([literarydevices.net](http://literarydevices.net))

Old farmer had a pig, a dog, a cow and a horse.

Old farmer had a pig, a dog, a cow and a horse.
**Balance: Enumeration**

**Enumeration** - a rhetorical device used to list a series of details, words or phrases. (literarydevices.net)

Old farmer had a pig, a dog, a cow and a horse.

Old farmer had a pig, a dog, a cow and a horse.
Balance: Enumeration

Enumeration - a rhetorical device used to list a series of details, words or phrases. (literarydevices.net)

Old farmer had a pig, a dog, a cow and a horse.

Old farmer had a pig, a dog, a cow and a horse.
Balance: Enumeration

**Enumeration** - a rhetorical device used to list a series of details, words or phrases. ([literarydevices.net](http://literarydevices.net))

```
Old farmer had a pig, a dog, a cow and a horse.
```

**UIMA Ruta**

```
Old farmer had a pig, a dog, a cow and a horse.
```
Balance: Enumeration

**Enumeration** - a rhetorical device used to list a series of details, words or phrases. (literarydevices.net)

Old farmer had a pig, a dog, a cow and a horse.

UIMA Ruta

Old farmer had a pig, a dog, a cow and a horse.
Balance: Enumeration

 Enumeration - a rhetorical device used to list a series of details, words or phrases. (literarydevices.net)

Old farmer had a pig, a dog, a cow and a horse.

UIMA Ruta

Old farmer had a pig, a dog, a cow and a horse.
Balance: Enumeration

**Enumeration** - a rhetorical device used to list a series of details, words or phrases. ([literarydevices.net](http://literarydevices.net))

Old farmer had a pig, a dog, a cow and a horse.

[UIMA Ruta]

Old farmer had a pig, a dog, a cow and a horse.
Balance: Enumeration

**Enumeration** - a rhetorical device used to list a series of details, words or phrases. (literarydevices.net)

Old farmer had a pig, a dog, a cow and a horse.

UIMA Ruta

Old farmer had a pig, a dog, a cow and a horse.
Balance: Enumeration

**Enumeration** - a rhetorical device used to list a series of details, words or phrases. ([literarydevices.net](http://literarydevices.net))

Old farmer had a pig, a dog, a cow and a horse.

Old farmer had a pig, a dog, a cow and a horse.
Enumeration - a rhetorical device used to list a series of details, words or phrases. (literarydevices.net)

Old farmer had a pig, a dog, a cow and a horse.

Old farmer had a pig, a dog, a cow and a horse.
Balance: Enumeration

Old farmer had a pig, a dog, a cow and a horse.

UIMA Ruta

Old farmer had a pig, a dog, a cow and a horse.
Balance: Enumeration

**Enumeration** - a rhetorical device used to list a series of details, words or phrases. (literarydevices.net)

Old farmer had a pig, a dog, a cow and a horse.

Old farmer had a pig, a dog, a cow and a horse.
Omission: Hypozeugma

Hypozeugma - placing last, in a construction containing several words or phrases of equal value, the word or words on which all of them depend. (Silva Rhetoricae)

A rooster, a prince and a lion walk into a bar...
Omission: Hypozeugma

Hypozeugma - placing last, in a construction containing several words or phrases of equal value, the word or words on which all of them depend. (Silva Rhetoricae)
Omission: Hypozeugma

**Hypozeugma** - placing last, in a construction containing several words or phrases of equal value, the word or words on which all of them depend. (Silva Rhetoricae)
Omission: Hypozeugma

Hypozeugma - placing last, in a construction containing several words or phrases of equal value, the word or words on which all of them depend. (Silva Rhetoricae)

A rooster, a prince and a lion walk into a bar…

Stanford Dependencies

A rooster, a prince and a lion walk into a bar…

governor-dependent relation
Omission: Hypozeugma

Hypozeugma - placing last, in a construction containing several words or phrases of equal value, the word or words on which all of them depend. (Silva Rhetoricae)
Omission: Hypozeugma

Hypozeugma - placing last, in a construction containing several words or phrases of equal value, the word or words on which all of them depend. *(Silva Rhetoricae)*

A rooster, a prince and a lion walk into a bar...

Stanford Dependencies

A rooster, a prince and a lion walk into a bar...

UIMA Ruta

rooster, a prince and a lion walk
Omission: Hypozeugma

Hypozeugma - placing last, in a construction containing several words or phrases of equal value, the word or words on which all of them depend. (Silva Rhetoricae)
Omission: Hypozeugma

Hypozeugma - placing last, in a construction containing several words or phrases of equal value, the word or words on which all of them depend. (Silva Rhetoricae)

A rooster, a prince and a lion walk into a bar…

Stanford Dependencies

A rooster, a prince and a lion walk into a bar…

UIMA Ruta

rooster, a prince and a lion walk
Omission: Hypozeugma

**Hypozeugma** - placing last, in a construction containing several words or phrases of equal value, the word or words on which all of them depend. (Silva Rhetoricae)

A rooster, a prince and a lion walk into a bar...

Stanford Dependencies

A rooster, a prince and a lion walk into a bar...

UIMA Ruta

rooster, a prince and a lion walk
Omission: Hypozeugma

Hypozeugma - placing last, in a construction containing several words or phrases of equal value, the word or words on which all of them depend. (Silva Rhetoricae)
**Omission: Hypozeugma**

**Hypozeugma** - placing last, in a construction containing several words or phrases of equal value, the word or words on which all of them depend. *(Silva Rhetoricae)*

Stanford Dependencies

A rooster, a prince and a lion walk into a bar...

UIMA Ruta

rooster, a prince and a lion walk
Repetition: Epanalepsis

Epanalepsis - repeats the beginning word of a sentence at the end.

Our eyes saw it, but we could not believe our eyes.
Repetition: Epanalepsis

Epanalepsis - repeats the beginning word of a sentence at the end.

Our eyes saw it, but we could not believe our eyes.
Repetition: Epanalepsis

Epanalepsis - repeats the beginning word of a sentence at the end.

Our eyes saw it, but we could not believe our eyes.

Our eyes saw believe our eyes.
Repetition: Epanalepsis

Epanalepsis - repeats the beginning word of a sentence at the end.

Our eyes saw it, but we could not believe our eyes.

Our eyes saw believe our eyes.
Repetition: Epanalepsis

_Epanalepsis_ - repeats the beginning word of a sentence at the end.

Our eyes saw it, but we could not believe our eyes.

Our eyes saw believe our eyes.
Repetition: Epanalepsis

Epanalepsis - repeats the beginning word of a sentence at the end.

Our eyes saw it, but we could not believe our eyes.

Our eyes saw believe our eyes.
Repetition: Epanalepsis

Epanalepsis - repeats the beginning word of a sentence at the end.

Our eyes saw it, but we could not believe our eyes.

Our eyes saw *believe* our eyes.
Repetition: Epanalepsis

Epanalepsis - repeats the beginning word of a sentence at the end.

Our eyes saw it, but we could not believe our eyes.

Our eyes saw believe our eyes.
Repetition: Epanalepsis

Epanalepsis - repeats the beginning word of a sentence at the end.

Our eyes saw it, but we could not believe our eyes.

Our eyes saw believe our eyes.
Custom: If-conditional 2

If-conditional 2 - expresses consequences that are totally unrealistic or will not likely happen in the future.

If I were president, I would cut taxes.
Custom: If-conditional 2

*If-conditional 2* - expresses consequences that are totally unrealistic or will not likely happen in the future.

If I were president, I would cut taxes.

Stanford Dependencies

If I were president, I would cut taxes.
Custom: If-conditional 2

If-conditional 2 - expresses consequences that are totally unrealistic or will not likely happen in the future.

If I were president, I would cut taxes.
If-conditional 2 - expresses consequences that are totally unrealistic or will not likely happen in the future.

If I were president, I would cut taxes.

Stanford Dependencies

If I were president, I would cut taxes.
Custom: If-conditional 2

If-conditional 2 - expresses consequences that are totally unrealistic or will not likely happen in the future.

If I were president, I would cut taxes.
Custom: If-conditional 2

*If-conditional 2* - expresses consequences that are totally unrealistic or will not likely happen in the future.

If I were president, I would cut taxes.
If-conditional 2 - expresses consequences that are totally unrealistic or will not likely happen in the future.

If I were president, I would cut taxes.
Custom: If-conditional 2

If-conditional 2 - expresses consequences that are totally unrealistic or will not likely happen in the future.

If I were president, I would cut taxes.

Stanford Dependencies

If I were president, I would cut taxes.

UIMA Ruta

If I were president I would cut
Custom: If-conditional 2

If-conditional 2 - expresses consequences that are totally unrealistic or will not likely happen in the future.

If I were president, I would cut taxes.

Stanford Dependencies

If I were president, I would cut taxes.

UIMA Ruta

If I were president, I would cut
Custom: If-conditional 2

If-conditional 2 - expresses consequences that are totally unrealistic or will not likely happen in the future.

If I were president, I would cut taxes.

Stanford Dependencies

If I were **president**, I would cut taxes.

UIMA Ruta

If I were **president** I **would** cut
Custom: If-conditional 2

**If-conditional 2** - expresses consequences that are totally unrealistic or will not likely happen in the future.

If I were president, I would cut taxes.
Evaluation dataset

- Expert websites
- Commercials
- Literature
- Political speeches
- Bible

Rhetorical Devices
Evaluation dataset

Evaluation measures

\[
\text{Precision} = \frac{tp}{tp + fp} \quad \text{Recall} = \frac{tp}{tp + fn} \quad F1 \text{ score} = 2 \cdot \frac{\text{precision} \cdot \text{recall}}{\text{precision} + \text{recall}}
\]
Evaluation Results

Balance schemes

- Enumeration: 0.84
- Isocolon: 0.68
- Pysma: 1

Omission schemes

- Asynd.: 0.4
- Epizeugma: 0.69
- Hypozeugma: 0.69
Evaluation Results

Balance schemes

- Enumeration: 0.84
- Isocolon: 0.68
- Pysma: 1

Omission schemes

- Asynd.: 0.4
- Epizeugma: 0.69
- Hypozeugma: 0.69

Precision
Recall
Evaluation Results

Old farmer had a pig, a dog, a cow and a horse.
Evaluation Results

Balance schemes

- Enumeration: 0.84
- Isocolon: 0.68
- Pysma: 1

Repetition schemes

- Anadiplosis: 0.74
- Diacope: 0.74
- Epanalepsis: 0.72
- Epiphoza: 0.74
- Epizeuxis: 0.78
- Mesarchia: 0.59
- Mesod: 0.4
- Polysyndeton: 0.73

Omission schemes

- Asynd.: 0.4
- Epizeugma: 0.69
- Hypozeugma: 0.69

Custom schemes

- If-cond 0: 0.73
- If-cond 1: 0.78
- If-cond 2: 0.78
- If-cond 3: 0.74
- If-counterFact: 0.85
- Passed Voice: 0.87
- Comp. Adj.: 0.56
- Comp. Adv.: 0.61
- Super. Adj.: 0.67
- Super. Adj.: 0.56
- Unless-cond: 1
- Whether-cond: 0.91
Evaluation Results

Repetition schemes
- Anadiplosis: 0.74
- Diacope: 0.74
- Epanalepsis: 0.72
- Epiphoza: 0.74
- Epizeuxis: 0.78
- Mesarchia: 0.59
- Mesod: 0.4
- Polysyndeton: 0.73

Balance schemes
- Enumeration: 0.84
- Isocolon: 0.68
- Pysma: 1

Omission schemes
- Asynd.: 0.4
- Epizeugma: 0.69
- Hypozeugma: 0.69

Custom schemes
- If-cond 0: 0.73
- If-cond 1: 0.78
- If-cond 2: 0.78
- If-cond 3: 0.78
- If-counterfact: 0.74
- If-counterfact: 0.85
- If-counterfact: 0.87
- Passive Voice: 0.56
- Comp. Adv.: 0.61
- Sup. Adv.: 0.67
- Unless-cond: 0.56
- Whether-cond: 1
- Unless-cond: 0.91
Evaluation Results F1-Score

Repetition schemes

- Anadiplosis: 0.74
- Diacope: 0.74
- Epanalepsis: 0.72
- Epiphora: 0.74
- Epizeuxis: 0.78
- Mesarchia: 0.59
- Mesod: 0.4
- Polysyndeton: 0.73

Balance schemes

- Enumeration: 0.84
- Isocolon: 0.68
- Pysma: 1

Omission schemes

- Asynd.: 0.4
- Epizeugma: 0.69
- Hypozeugma: 0.69

Custom schemes

- If-cond 0: 0.73
- If-cond 1: 0.78
- If-cond 2: 0.78
- If-cond 3: 0.74
- If-counterfact: 0.85
- Passive Voice: 0.87
- Comp. Adv.: 0.56
- Super. Adv.: 0.61
- Unless-cond: 0.67
- Whether-cond: 1

Balance schemes

- Enumeration: 0.84
- Isocolon: 0.68
- Pysma: 1

Omission schemes

- Asynd.: 0.4
- Epizeugma: 0.69
- Hypozeugma: 0.69

Custom schemes

- If-cond 0: 0.73
- If-cond 1: 0.78
- If-cond 2: 0.78
- If-cond 3: 0.74
- If-counterfact: 0.85
- Passive Voice: 0.87
- Comp. Adv.: 0.56
- Super. Adv.: 0.61
- Unless-cond: 0.67
- Whether-cond: 1
Pipeline
Analysis of Rhetorical Devices
Pipeline
Pipeline

Frequency

Combinations

Group distribution
Pipeline

![Diagram of pipeline with output, distribution statistics, and confidence interval](image)

Confidence interval:
- \( n \times \text{precision (rd)} \)
- \( n / \text{recall (rd)} \)

Device:
- Precision
- Recall

Al-Khatib et al. [2017]
Pipeline

Distribution Statistics

Significance Test

Effect-size Test

Misclassification Impact

output

Significance Test

A

B

A

B
Pipeline
Pipeline
3 Analysis Experiments

Data Preparation
Experiments: datasets

The New York Times

US Presidential Debates 2016

Ben Wiseman [2016]
Data dimensionality

<table>
<thead>
<tr>
<th>Language</th>
<th>Mode</th>
<th>Communication</th>
<th>Author</th>
<th>Audience</th>
</tr>
</thead>
<tbody>
<tr>
<td>English</td>
<td>Written</td>
<td>Monological</td>
<td>Identity</td>
<td>U.S.</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Type</th>
<th>Genre</th>
<th>Topic</th>
<th>Medium</th>
</tr>
</thead>
<tbody>
<tr>
<td>Descriptive</td>
<td>Editorial</td>
<td>Education</td>
<td>Newspaper</td>
</tr>
<tr>
<td>Argumentative</td>
<td>Review</td>
<td>Science</td>
<td>Presidential Debates</td>
</tr>
<tr>
<td>Biography</td>
<td></td>
<td>Art</td>
<td></td>
</tr>
<tr>
<td>Debate</td>
<td></td>
<td>Politics</td>
<td></td>
</tr>
</tbody>
</table>
NYT Experiment: data subsampling

NYT Dataset

Random
1000 articles

Article-length based
600 articles

Matching
343 articles
NYT Experiment: Findings

“Random” dataset

Rhetoric Detection System

“Article-length based” dataset

Rhetoric Detection System

Articles cover multiple dimensions

Hard to deduce particular styles
NYT Experiment: Findings

“Random” dataset

“Article-length based” dataset

Rhetoric Detection System

CONFOUNDING

Rhetoric Detection System

Articles cover multiple dimensions

Hard to deduce particular styles
NYT Experiment: Confounding

Genre
(confounding variable)

Author
(independent variable)

Rhetorical style
(dependent variable)
NYT Experiment: Confounding

Genre
(confounding variable)

Author
(independent variable)

Rhetorical style
(dependent variable)

MATCHING
NYT Experiment: Matching
NYT Experiment: Matching
NYT Experiment: Matching
NYT Experiment: Matching

Articles written by author X on topic Y across genres
NYT Experiment: Matching
NYT Experiment: Matching
NYT Experiment: Matching
NYT Experiment: Matching
NYT Experiment: Matching
Analysis
Experiments

Findings
NYT Experiment: Findings

Style-based frequency of rhetorical devices
NYT Experiment: Findings

<table>
<thead>
<tr>
<th>Author</th>
<th>EPIPHOZA Distribution (%)</th>
<th>REPETITION SCHEMES Distribution (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Hevesi Dennis</td>
<td>10.74</td>
<td>70.99</td>
</tr>
<tr>
<td>Lewis Paul</td>
<td>12.99</td>
<td>81.93</td>
</tr>
<tr>
<td>Martin Douglas</td>
<td>6.49</td>
<td>55.49</td>
</tr>
</tbody>
</table>
NYT Experiment: Findings

<table>
<thead>
<tr>
<th>Author</th>
<th>EPIPHOZA Distribution (%)</th>
<th>REPETITION SCHEMES Distribution (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Hevesi Dennis</td>
<td>10.74</td>
<td>70.99</td>
</tr>
<tr>
<td>Lewis Paul</td>
<td>12.99</td>
<td>81.93</td>
</tr>
<tr>
<td>Martin Douglas</td>
<td>6.49</td>
<td>55.49</td>
</tr>
</tbody>
</table>

! Same pattern across all articles
NYT Experiment: Findings

<table>
<thead>
<tr>
<th>Author</th>
<th>Distribution (%)</th>
<th>EPIPHOZA</th>
<th>Distribution (%)</th>
<th>REPETITION SCHEMES</th>
</tr>
</thead>
<tbody>
<tr>
<td>Hevesi Dennis</td>
<td>10.74</td>
<td></td>
<td>70.99</td>
<td></td>
</tr>
<tr>
<td>Lewis Paul</td>
<td>12.99</td>
<td></td>
<td>81.93</td>
<td></td>
</tr>
<tr>
<td>Martin Douglas</td>
<td>6.49</td>
<td></td>
<td>55.49</td>
<td></td>
</tr>
</tbody>
</table>

Good Job, Lewis!
## NYT Experiment: Findings

### Authors

<table>
<thead>
<tr>
<th>Datasets</th>
<th>P-value</th>
<th>Independence</th>
<th>Cramer’s V value</th>
<th>Effect</th>
</tr>
</thead>
<tbody>
<tr>
<td>Hevesi vs. Lewis</td>
<td>0.015</td>
<td>TRUE*</td>
<td>0.1</td>
<td>SMALL</td>
</tr>
<tr>
<td>Lewis vs. Martin</td>
<td>~0</td>
<td>TRUE</td>
<td>0.15</td>
<td>SMALL</td>
</tr>
<tr>
<td>Martin vs. Hevesi</td>
<td>0.017</td>
<td>TRUE*</td>
<td>0.1</td>
<td>SMALL</td>
</tr>
</tbody>
</table>

* for $\alpha > 0.001$
NYT Experiment: Findings

Genres

<table>
<thead>
<tr>
<th>Confounders</th>
<th>Distribution (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Genre:</td>
<td>Biography</td>
</tr>
<tr>
<td>freedman-news</td>
<td>11.65</td>
</tr>
<tr>
<td>norris-markets</td>
<td>22.59</td>
</tr>
<tr>
<td>wade-health</td>
<td>12.04</td>
</tr>
</tbody>
</table>

author  topic
### NYT Experiment: Findings

<table>
<thead>
<tr>
<th>Genres</th>
<th>COMPARATIVES</th>
<th>CONDITIONALS</th>
</tr>
</thead>
<tbody>
<tr>
<td>Genre</td>
<td>Distribution (%)</td>
<td>Distribution (%)</td>
</tr>
<tr>
<td>Biography</td>
<td>14.07</td>
<td>3.45</td>
</tr>
<tr>
<td>Editorial</td>
<td>23.16</td>
<td>5.95</td>
</tr>
<tr>
<td>Review</td>
<td>16.29</td>
<td>3.41</td>
</tr>
</tbody>
</table>
NYT Experiment: Findings

<table>
<thead>
<tr>
<th>Datasets</th>
<th>SIGNIFICANCE</th>
<th>EFFECT-SIZE</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>P-value</td>
<td>Cramer’s V value</td>
</tr>
<tr>
<td>Biography vs. Editorial</td>
<td>~0, TRUE</td>
<td>0.16</td>
</tr>
<tr>
<td>Editorial vs. Review</td>
<td>~0, TRUE</td>
<td>0.14</td>
</tr>
<tr>
<td>Review vs. Biography</td>
<td>0.68, FALSE</td>
<td>0.07</td>
</tr>
</tbody>
</table>
NYT Experiment: Findings

<table>
<thead>
<tr>
<th>Confounders</th>
<th>Distribution (%)</th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Topics:</td>
<td></td>
<td>Arts</td>
<td>Education</td>
<td>Science</td>
</tr>
<tr>
<td>martin-biography</td>
<td>11.95</td>
<td>10.94</td>
<td>12.24</td>
<td></td>
</tr>
<tr>
<td>saxon-biography</td>
<td>6.15</td>
<td>6.14</td>
<td>12.50</td>
<td></td>
</tr>
</tbody>
</table>
NYT Experiment: Findings

Style-based frequency of rhetorical devices

Characteristic style patterns within each dimension
NYT Experiment: Findings

<table>
<thead>
<tr>
<th>Datasets</th>
<th>P-value</th>
<th>Independence</th>
<th>Cramer’s V value</th>
<th>Effect</th>
</tr>
</thead>
<tbody>
<tr>
<td>Science vs. Education</td>
<td>0.70</td>
<td>FALSE</td>
<td>0.09</td>
<td>SMALL</td>
</tr>
<tr>
<td>Education vs. Arts</td>
<td>0.26</td>
<td>FALSE</td>
<td>0.10</td>
<td>SMALL</td>
</tr>
<tr>
<td>Arts vs. Science</td>
<td>0.19</td>
<td>FALSE</td>
<td>0.10</td>
<td>SMALL</td>
</tr>
</tbody>
</table>
NYT Experiment: Findings

- Style-based frequency of rhetorical devices
- Characteristic style patterns within each dimension
- Style is more author- and genre-dependent
Presidential Debates: Datasets

CLINTON VS. TRUMP

CLINTON VS. REST

TRUMP VS. CLINTON

TRUMP VS. REST
Presidential Debates: Findings

<table>
<thead>
<tr>
<th>Debate Type</th>
<th>ASYNDETON Distribution (%)</th>
<th>VOICE Distribution (%)</th>
<th>BALANCE SCH. Distribution (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Clinton → Trump</td>
<td>15.24</td>
<td>8.07</td>
<td>17.69</td>
</tr>
<tr>
<td>Trump → Clinton</td>
<td>10.83</td>
<td>5.29</td>
<td>19.92</td>
</tr>
</tbody>
</table>
Presidential Debates: Findings

<table>
<thead>
<tr>
<th>Debate Type</th>
<th>ASYNDETON Distribution (%)</th>
<th>VOICE Distribution (%)</th>
<th>BALANCE SCH. Distribution (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Clinton → Trump</td>
<td>15.24</td>
<td>8.07</td>
<td>17.69</td>
</tr>
<tr>
<td>Trump → Clinton</td>
<td>10.83</td>
<td>5.29</td>
<td>19.92</td>
</tr>
</tbody>
</table>

Asyndeton = clarity and rhythm
Presidential Debates: Findings

<table>
<thead>
<tr>
<th>Debate Type</th>
<th>Distribution (%)</th>
<th>Distribution (%)</th>
<th>Distribution (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Clinton → Trump</td>
<td>15.24</td>
<td>8.07</td>
<td>17.69</td>
</tr>
<tr>
<td>Trump → Clinton</td>
<td>10.83</td>
<td>5.29</td>
<td>19.92</td>
</tr>
</tbody>
</table>

Acceptance Speech Analysis by Huffington Post

<table>
<thead>
<tr>
<th>Candidate</th>
<th>Sent.</th>
<th>Long Sent. (%)</th>
<th>Passive voice (%)</th>
<th>Grade Level (US)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Hillary Clinton</td>
<td>413</td>
<td>7.26</td>
<td>3.39</td>
<td>5</td>
</tr>
<tr>
<td>Donald Trump</td>
<td>341</td>
<td>16.42</td>
<td>8.8</td>
<td>8</td>
</tr>
</tbody>
</table>
# Presidential Debates: Findings

## Analysis by Huffington Post

<table>
<thead>
<tr>
<th>Debate Type</th>
<th>ASYNDETON</th>
<th>VOICE</th>
<th>BALANCE SCH.</th>
</tr>
</thead>
<tbody>
<tr>
<td>Clinton → Trump</td>
<td>15.24</td>
<td>8.07</td>
<td>17.69</td>
</tr>
<tr>
<td>Trump → Clinton</td>
<td>10.83</td>
<td>5.29</td>
<td>19.92</td>
</tr>
</tbody>
</table>

### Acceptance Speech Analysis by Huffington Post

<table>
<thead>
<tr>
<th>Candidate</th>
<th>Sent.</th>
<th>Long Sent. (%)</th>
<th>Passive voice (%)</th>
<th>Grade Level (US)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Hillary Clinton</td>
<td>413</td>
<td>7.26</td>
<td>3.39</td>
<td>5</td>
</tr>
<tr>
<td>Donald Trump</td>
<td>341</td>
<td>16.42</td>
<td>8.8</td>
<td>8</td>
</tr>
</tbody>
</table>
## Presidential Debates: Findings

<table>
<thead>
<tr>
<th>Debate Type</th>
<th>ASYNDETON</th>
<th>VOICE</th>
<th>BALANCE SCH.</th>
</tr>
</thead>
<tbody>
<tr>
<td>Clinton → Trump</td>
<td>15.24</td>
<td>8.07</td>
<td>17.69</td>
</tr>
<tr>
<td>Trump → Clinton</td>
<td>10.83</td>
<td>5.29</td>
<td>19.92</td>
</tr>
</tbody>
</table>

### Acceptance Speech Analysis by Huffington Post

<table>
<thead>
<tr>
<th>Candidate</th>
<th>Sent.</th>
<th>Long Sent. (%)</th>
<th>Passive voice (%)</th>
<th>Grade Level (US)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Hillary Clinton</td>
<td>413</td>
<td>7.26</td>
<td>3.39</td>
<td>5</td>
</tr>
<tr>
<td>Donald Trump</td>
<td>341</td>
<td>16.42</td>
<td>8.8</td>
<td>8</td>
</tr>
</tbody>
</table>
### Presidential Debates: Findings

#### Significance Test

<table>
<thead>
<tr>
<th>Debate Type</th>
<th>Clinton $\rightarrow$ Rest</th>
<th>Clinton $\rightarrow$ Trump</th>
<th>Trump $\rightarrow$ Clinton</th>
<th>Trump $\rightarrow$ Rest</th>
</tr>
</thead>
<tbody>
<tr>
<td>Clinton $\rightarrow$ Rest</td>
<td>TRUE*</td>
<td>TRUE*</td>
<td>TRUE</td>
<td>TRUE</td>
</tr>
<tr>
<td>Clinton $\rightarrow$ Trump</td>
<td>TRUE*</td>
<td>TRUE</td>
<td>TRUE</td>
<td>TRUE</td>
</tr>
<tr>
<td>Trump $\rightarrow$ Clinton</td>
<td>TRUE</td>
<td>TRUE</td>
<td>TRUE*</td>
<td>FALSE†</td>
</tr>
<tr>
<td>Trump $\rightarrow$ Rest</td>
<td>TRUE</td>
<td>TRUE</td>
<td>FALSE†</td>
<td>TRUE</td>
</tr>
</tbody>
</table>

* for $\alpha > 0.01$

† for $\alpha > 0.1$
Presidential Debates: Findings

Significance Test

<table>
<thead>
<tr>
<th>Debate Type</th>
<th>Clinton → Rest</th>
<th>Clinton → Trump</th>
<th>Trump → Clinton</th>
<th>Trump → Rest</th>
</tr>
</thead>
<tbody>
<tr>
<td>Clinton → Rest</td>
<td>TRUE*</td>
<td>TRUE*</td>
<td>TRUE</td>
<td>TRUE</td>
</tr>
<tr>
<td>Clinton → Trump</td>
<td>TRUE*</td>
<td>TRUE</td>
<td>TRUE</td>
<td>TRUE</td>
</tr>
<tr>
<td>Trump → Clinton</td>
<td>TRUE</td>
<td>TRUE</td>
<td>TRUE</td>
<td>FALSE†</td>
</tr>
<tr>
<td>Trump → Rest</td>
<td>TRUE</td>
<td>TRUE</td>
<td>TRUE</td>
<td>FALSE†</td>
</tr>
</tbody>
</table>

* for $\alpha > 0.01$
† for $\alpha > 0.1$

Trump doesn’t change his style
## Summary

### Conclusions

| System for rhetorical style identification in high-quality text documents |
| Rule-based algorithms for detection of RD |
| Vague style patterns across random and article-length based subsampling: **Confounding** |
| Better style identification with **Matching** |
| Rhetorical style depends more on author and genre of writings rather than their topics |
| Debates: candidates employ different styles |
| Debates: domain experience trains an adaptive rhetorical style |
Summary

Conclusions

System for rhetorical style identification in high-quality text documents

Rule-based algorithms for detection of RD

Vague style patterns across random and article-length based subsampling: Confounding

Better style identification with Matching

Rhetorical style depends more on author and genre of writings rather than their topics

Debates: candidates employ different styles

Debates: domain experience trains an adaptive rhetorical style

Resources

Novel framework for detecting rhetorical devices

Comprehensive dataset for evaluation of rhetoric detection systems

Elaborative style patterns and intriguing findings
Summary

Conclusions

System for rhetorical style identification in high-quality text documents
Rule-based algorithms for detection of RD
Vague style patterns across random and article-length based subsampling: Confounding
Better style identification with Matching
Rhetorical style depends more on author and genre of writings rather than their topics
Debates: candidates employ different styles
Debates: domain experience trains an adaptive rhetorical style

Resources

Novel framework for detecting rhetorical devices
Comprehensive dataset for evaluation of rhetoric detection systems
Elaborative style patterns and intriguing findings

Efficiency

1st sentence → 5.8 sec.
2nd sentence → 0.4 sec.
Initialization → 1.7 sec.
## Summary

### Conclusions

- System for rhetorical style identification in high-quality text documents
- Rule-based algorithms for detection of RD
- Vague style patterns across random and article-length based subsampling: **Confounding**
- Better style identification with **Matching**
- Rhetorical style depends more on author and genre of writings rather than their topics
- Debates: candidates employ different styles
- Debates: domain experience trains an adaptive rhetorical style

### Resources

- Novel framework for detecting rhetorical devices
- Comprehensive dataset for evaluation of rhetoric detection systems
- Elaborative style patterns and intriguing findings

### Efficiency

- 1\textsuperscript{st} sentence $\rightarrow$ 5.8 sec.
- 2\textsuperscript{nd} sentence $\rightarrow$ 0.4 sec.
- Initialization $\rightarrow$ 1.7 sec.

### Future Work

- Larger dataset for analysis
- Focus of semantical rhetoric
- Analysis measures like placement and flows of rhetorical devices
Thank you!
References


• https://literarydevices.net/enumeration/

• G. Burton. The forest of rhetoric (silva rhetoricae), 2007.
References - Icons and Images

- advertise by David from the Noun Project
- buy by Arthur Shlain from the Noun Project
- Money by Desbenoit from the Noun Project
- Idea by MRFA from the Noun Project
- arrange by Gregor Cresnar from the Noun Project
- font style by iconsmind.com from the Noun Project
- memories by Henning Gross from the Noun Project
- what by Paffi from the Noun Project
- Translation by Mun May Tee from the Noun Project
- analysis by Chameleon Design from the Noun Project
- like by Bluetip Design from the Noun Project
- analysis by Chameleon Design from the Noun Project
- Fokler by AlfredoCreates.com/Icons from the Noun Project
- different by AlfredoCreates.com/Icons from the Noun Project
- Flag by Hare Krishna from the Noun Project
- Map Marker by shashank singh from the Noun Project
- Icon by Liisole from the Noun Project
- Icons made by Freepik on flaticon.com
- Icons made by Becris on flaticon.com
- Icons made by Vectors Market on flaticon.com
- Jar by S. Salinas from the Noun Project
- Check mark designed by Freepik
- Icons made by Smashicons on flaticon.com
Existing research

• Gawryjołek et al. [2009] – authorship identification system based on rhetorical style.

• Strommer [2011] – authorial intent detection system based on the anaphora usage.

• Java [2015] – machine-learning based authorship identification system using rhetorical devices (based on Gawryjołek et al. [2009]).
## Evaluation results

<table>
<thead>
<tr>
<th>Device</th>
<th>Total No.</th>
<th>Precision</th>
<th>Recall</th>
<th>F1-score</th>
<th>Device</th>
<th>Total No.</th>
<th>Precision</th>
<th>Recall</th>
<th>F1-score</th>
</tr>
</thead>
<tbody>
<tr>
<td>Anadiplosis</td>
<td>60</td>
<td>0.76</td>
<td>0.73</td>
<td>0.74</td>
<td>If Conditional Two</td>
<td>60</td>
<td>0.82</td>
<td>0.75</td>
<td>0.78</td>
</tr>
<tr>
<td>Asyndeton</td>
<td>60</td>
<td>0.25</td>
<td>0.95</td>
<td>0.4</td>
<td>If Conditional Zero</td>
<td>60</td>
<td>0.71</td>
<td>0.76</td>
<td>0.73</td>
</tr>
<tr>
<td>Comparative Adjective</td>
<td>67</td>
<td>0.51</td>
<td>0.61</td>
<td>0.56</td>
<td>If Counterfactual</td>
<td>60</td>
<td>0.84</td>
<td>0.87</td>
<td>0.85</td>
</tr>
<tr>
<td>Comparative Adverb</td>
<td>71</td>
<td>0.6</td>
<td>0.62</td>
<td>0.61</td>
<td>Isocolon</td>
<td>180</td>
<td>0.57</td>
<td>0.83</td>
<td>0.68</td>
</tr>
<tr>
<td>Diacope</td>
<td>60</td>
<td>0.75</td>
<td>0.73</td>
<td>0.74</td>
<td>Mesarchia</td>
<td>20</td>
<td>0.45</td>
<td>0.85</td>
<td>0.59</td>
</tr>
<tr>
<td>Enumeration</td>
<td>60</td>
<td>0.76</td>
<td>0.93</td>
<td>0.84</td>
<td>Mesodiplosis</td>
<td>40</td>
<td>0.28</td>
<td>0.68</td>
<td>0.4</td>
</tr>
<tr>
<td>Epanalepsis</td>
<td>60</td>
<td>0.63</td>
<td>0.83</td>
<td>0.72</td>
<td>Passive Voice</td>
<td>60</td>
<td>0.79</td>
<td>0.98</td>
<td>0.87</td>
</tr>
<tr>
<td>Epiphoza</td>
<td>60</td>
<td>0.61</td>
<td>0.93</td>
<td>0.74</td>
<td>Polysyndeton</td>
<td>60</td>
<td>0.77</td>
<td>0.7</td>
<td>0.73</td>
</tr>
<tr>
<td>Epizeugma</td>
<td>60</td>
<td>0.68</td>
<td>0.7</td>
<td>0.69</td>
<td>Pysma</td>
<td>60</td>
<td>1</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>Epizeuxis</td>
<td>60</td>
<td>0.79</td>
<td>0.77</td>
<td>0.78</td>
<td>Superlative Adjective</td>
<td>70</td>
<td>0.62</td>
<td>0.73</td>
<td>0.67</td>
</tr>
<tr>
<td>Hypozeugma</td>
<td>60</td>
<td>0.61</td>
<td>0.8</td>
<td>0.69</td>
<td>Superlative Adverb</td>
<td>70</td>
<td>0.63</td>
<td>0.5</td>
<td>0.56</td>
</tr>
<tr>
<td>If Conditional One</td>
<td>60</td>
<td>0.78</td>
<td>0.78</td>
<td>0.78</td>
<td>Unless Conditional</td>
<td>60</td>
<td>1</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>If Conditional Three</td>
<td>60</td>
<td>0.86</td>
<td>0.65</td>
<td>0.74</td>
<td>Whether Conditional</td>
<td>60</td>
<td>1</td>
<td>0.83</td>
<td>0.91</td>
</tr>
</tbody>
</table>

- **Balance schemes**
- **Omission schemes**
- **Repetition schemes**
- **Custom schemes**
### Evaluation results

<table>
<thead>
<tr>
<th>Device</th>
<th>Total No.</th>
<th>Precision</th>
<th>Recall</th>
<th>F1-score</th>
</tr>
</thead>
<tbody>
<tr>
<td>Anadiplosis</td>
<td>60</td>
<td>0.76</td>
<td>0.73</td>
<td>0.74</td>
</tr>
<tr>
<td>Asyndeton</td>
<td>60</td>
<td>0.25</td>
<td>0.95</td>
<td>0.4</td>
</tr>
<tr>
<td>Comparative Adjective</td>
<td>67</td>
<td>0.51</td>
<td>0.61</td>
<td>0.56</td>
</tr>
<tr>
<td>Comparative Adverb</td>
<td>71</td>
<td>0.6</td>
<td>0.62</td>
<td>0.61</td>
</tr>
<tr>
<td>Diacope</td>
<td>60</td>
<td>0.75</td>
<td>0.73</td>
<td>0.74</td>
</tr>
<tr>
<td>Enumeration</td>
<td>60</td>
<td>0.76</td>
<td>0.93</td>
<td>0.84</td>
</tr>
<tr>
<td>Epanalepsis</td>
<td>60</td>
<td>0.63</td>
<td>0.83</td>
<td>0.72</td>
</tr>
<tr>
<td>Epiphoza</td>
<td>60</td>
<td>0.61</td>
<td>0.93</td>
<td>0.74</td>
</tr>
<tr>
<td>Epizeugma</td>
<td>60</td>
<td>0.68</td>
<td>0.7</td>
<td>0.69</td>
</tr>
<tr>
<td>Epizeuxis</td>
<td>60</td>
<td>0.79</td>
<td>0.77</td>
<td>0.78</td>
</tr>
<tr>
<td>Hypozeugma</td>
<td>60</td>
<td>0.61</td>
<td>0.8</td>
<td>0.69</td>
</tr>
<tr>
<td>If Conditional One</td>
<td>60</td>
<td>0.78</td>
<td>0.78</td>
<td>0.78</td>
</tr>
<tr>
<td>If Conditional Three</td>
<td>60</td>
<td>0.86</td>
<td>0.65</td>
<td>0.74</td>
</tr>
<tr>
<td>If Conditional Two</td>
<td>60</td>
<td>0.82</td>
<td>0.75</td>
<td>0.78</td>
</tr>
<tr>
<td>If Conditional Zero</td>
<td>60</td>
<td>0.71</td>
<td>0.76</td>
<td>0.73</td>
</tr>
<tr>
<td>If Counterfactual</td>
<td>60</td>
<td>0.84</td>
<td>0.87</td>
<td>0.85</td>
</tr>
<tr>
<td>Isocolon</td>
<td>180</td>
<td>0.57</td>
<td>0.83</td>
<td>0.68</td>
</tr>
<tr>
<td>Mesarchia</td>
<td>20</td>
<td>0.45</td>
<td>0.85</td>
<td>0.59</td>
</tr>
<tr>
<td>Mesodiplosis</td>
<td>40</td>
<td>0.28</td>
<td>0.68</td>
<td>0.4</td>
</tr>
<tr>
<td>Passive Voice</td>
<td>60</td>
<td>0.79</td>
<td>0.98</td>
<td>0.87</td>
</tr>
<tr>
<td>Polysyndeton</td>
<td>60</td>
<td>0.77</td>
<td>0.7</td>
<td>0.73</td>
</tr>
<tr>
<td>Pysma</td>
<td>60</td>
<td>1</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>Superlative Adjective</td>
<td>70</td>
<td>0.62</td>
<td>0.73</td>
<td>0.67</td>
</tr>
<tr>
<td>Superlative Adverb</td>
<td>70</td>
<td>0.63</td>
<td>0.5</td>
<td>0.56</td>
</tr>
<tr>
<td>Unless Conditional</td>
<td>60</td>
<td>1</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>Whether Conditional</td>
<td>60</td>
<td>1</td>
<td>0.83</td>
<td>0.91</td>
</tr>
</tbody>
</table>

- **Balance schemes**
- **Omission schemes**
- **Repetition schemes**
- **Custom schemes**
## Evaluation results

<table>
<thead>
<tr>
<th>Device</th>
<th>Total No.</th>
<th>Precision</th>
<th>Recall</th>
<th>F1-score</th>
</tr>
</thead>
<tbody>
<tr>
<td>Anadiplosis</td>
<td>60</td>
<td>0.76</td>
<td>0.73</td>
<td>0.74</td>
</tr>
<tr>
<td>Asyndeton</td>
<td>60</td>
<td>0.25</td>
<td>0.95</td>
<td>0.40</td>
</tr>
<tr>
<td><strong>Comparative Adjective</strong></td>
<td>67</td>
<td>0.51</td>
<td>0.61</td>
<td>0.56</td>
</tr>
<tr>
<td><strong>Comparative Adverb</strong></td>
<td>71</td>
<td>0.60</td>
<td>0.62</td>
<td>0.61</td>
</tr>
<tr>
<td>Diacope</td>
<td>60</td>
<td>0.75</td>
<td>0.73</td>
<td>0.74</td>
</tr>
<tr>
<td>Enumeration</td>
<td>60</td>
<td>0.76</td>
<td>0.93</td>
<td>0.84</td>
</tr>
<tr>
<td>Epanalepsis</td>
<td>60</td>
<td>0.63</td>
<td>0.83</td>
<td>0.72</td>
</tr>
<tr>
<td>Epiphoza</td>
<td>60</td>
<td>0.61</td>
<td>0.93</td>
<td>0.74</td>
</tr>
<tr>
<td>Epizeugma</td>
<td>60</td>
<td>0.68</td>
<td>0.7</td>
<td>0.69</td>
</tr>
<tr>
<td>Epizeuxis</td>
<td>60</td>
<td>0.79</td>
<td>0.77</td>
<td>0.78</td>
</tr>
<tr>
<td>Hypozeugma</td>
<td>60</td>
<td>0.61</td>
<td>0.8</td>
<td>0.69</td>
</tr>
<tr>
<td>If Conditional One</td>
<td>60</td>
<td>0.78</td>
<td>0.78</td>
<td>0.78</td>
</tr>
<tr>
<td>If Conditional Three</td>
<td>60</td>
<td>0.86</td>
<td>0.65</td>
<td>0.74</td>
</tr>
<tr>
<td>If Conditional Two</td>
<td>60</td>
<td>0.82</td>
<td>0.75</td>
<td>0.78</td>
</tr>
<tr>
<td>If Conditional Zero</td>
<td>60</td>
<td>0.71</td>
<td>0.76</td>
<td>0.73</td>
</tr>
<tr>
<td>If Counterfactual</td>
<td>60</td>
<td>0.84</td>
<td>0.87</td>
<td>0.85</td>
</tr>
<tr>
<td>Isocolon</td>
<td>180</td>
<td>0.57</td>
<td>0.83</td>
<td>0.68</td>
</tr>
<tr>
<td>Mesarchia</td>
<td>20</td>
<td>0.45</td>
<td>0.85</td>
<td>0.59</td>
</tr>
<tr>
<td>Mesodiplosis</td>
<td>40</td>
<td>0.28</td>
<td>0.68</td>
<td>0.40</td>
</tr>
<tr>
<td>Passive Voice</td>
<td>60</td>
<td>0.79</td>
<td>0.98</td>
<td>0.87</td>
</tr>
<tr>
<td>Polysyndeton</td>
<td>60</td>
<td>0.77</td>
<td>0.7</td>
<td>0.73</td>
</tr>
<tr>
<td>Pysma</td>
<td>60</td>
<td>1.00</td>
<td>1.00</td>
<td>1.00</td>
</tr>
<tr>
<td>Superlative Adjective</td>
<td>70</td>
<td>0.62</td>
<td>0.73</td>
<td>0.67</td>
</tr>
<tr>
<td>Superlative Adverb</td>
<td>70</td>
<td>0.63</td>
<td>0.50</td>
<td>0.56</td>
</tr>
<tr>
<td>Unless Conditional</td>
<td>60</td>
<td>1.00</td>
<td>1.00</td>
<td>1.00</td>
</tr>
<tr>
<td>Whether Conditional</td>
<td>60</td>
<td>1.00</td>
<td>0.83</td>
<td>0.91</td>
</tr>
</tbody>
</table>

- Balance schemes
- Omission schemes
- Repetition schemes
- Custom schemes
## Evaluation results

<table>
<thead>
<tr>
<th>Device</th>
<th>Total No.</th>
<th>Precision</th>
<th>Recall</th>
<th>F1-score</th>
<th>Device</th>
<th>Total No.</th>
<th>Precision</th>
<th>Recall</th>
<th>F1-score</th>
<th>Device</th>
<th>Total No.</th>
<th>Precision</th>
<th>Recall</th>
<th>F1-score</th>
</tr>
</thead>
<tbody>
<tr>
<td>Anadiplosis</td>
<td>60</td>
<td>0.76</td>
<td>0.73</td>
<td>0.74</td>
<td>If Conditional Two</td>
<td>60</td>
<td>0.82</td>
<td>0.75</td>
<td>0.78</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Asyndeton</td>
<td>60</td>
<td>0.25</td>
<td>0.95</td>
<td>0.4</td>
<td>If Conditional Zero</td>
<td>60</td>
<td>0.71</td>
<td>0.76</td>
<td>0.73</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Comparative Adjective</td>
<td>67</td>
<td>0.51</td>
<td>0.61</td>
<td>0.56</td>
<td>If Counterfactual</td>
<td>60</td>
<td>0.84</td>
<td>0.87</td>
<td>0.85</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Comparative Adverb</td>
<td>71</td>
<td>0.6</td>
<td>0.62</td>
<td>0.61</td>
<td>Isocolon</td>
<td>180</td>
<td>0.57</td>
<td>0.83</td>
<td>0.68</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Diacope</td>
<td>60</td>
<td>0.75</td>
<td>0.73</td>
<td>0.74</td>
<td>Mesarchia</td>
<td>20</td>
<td>0.45</td>
<td>0.85</td>
<td>0.59</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Enumeration</td>
<td>60</td>
<td>0.76</td>
<td>0.93</td>
<td>0.84</td>
<td>Mesodiplosis</td>
<td>40</td>
<td>0.28</td>
<td>0.68</td>
<td>0.4</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Epanalepsis</td>
<td>60</td>
<td>0.63</td>
<td>0.83</td>
<td>0.72</td>
<td>Passive Voice</td>
<td>60</td>
<td>0.79</td>
<td>0.98</td>
<td>0.87</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Epiphoza</td>
<td>60</td>
<td>0.61</td>
<td>0.93</td>
<td>0.74</td>
<td>Polysyndeton</td>
<td>60</td>
<td>0.77</td>
<td>0.7</td>
<td>0.73</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Epizeugma</td>
<td>60</td>
<td>0.68</td>
<td>0.7</td>
<td>0.69</td>
<td>Pysma</td>
<td>60</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Epizeuxis</td>
<td>60</td>
<td>0.79</td>
<td>0.77</td>
<td>0.78</td>
<td>Superlative Adjective</td>
<td>70</td>
<td>0.62</td>
<td>0.73</td>
<td>0.67</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Hypozeugma</td>
<td>60</td>
<td>0.61</td>
<td>0.8</td>
<td>0.69</td>
<td>Superlative Adverb</td>
<td>70</td>
<td>0.63</td>
<td>0.5</td>
<td>0.56</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>If Conditional One</td>
<td>60</td>
<td>0.78</td>
<td>0.78</td>
<td>0.78</td>
<td>Unless Conditional</td>
<td>60</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>If Conditional Three</td>
<td>60</td>
<td>0.86</td>
<td>0.65</td>
<td>0.74</td>
<td>Whether Conditional</td>
<td>60</td>
<td>1</td>
<td>0.83</td>
<td>0.91</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

- Balance schemes
- Omission schemes
- Repetition schemes
- Custom schemes
# Evaluation results

<table>
<thead>
<tr>
<th>Device</th>
<th>Total No.</th>
<th>Precision</th>
<th>Recall</th>
<th>F1-score</th>
</tr>
</thead>
<tbody>
<tr>
<td>Anadiplosis</td>
<td>60</td>
<td>0.76</td>
<td>0.73</td>
<td>0.74</td>
</tr>
<tr>
<td>Asyndeton</td>
<td>60</td>
<td>0.25</td>
<td>0.95</td>
<td>0.4</td>
</tr>
<tr>
<td>Comparative Adjective</td>
<td>67</td>
<td>0.51</td>
<td>0.61</td>
<td>0.56</td>
</tr>
<tr>
<td>Comparative Adverb</td>
<td>71</td>
<td>0.6</td>
<td>0.62</td>
<td>0.61</td>
</tr>
<tr>
<td>Diacope</td>
<td>60</td>
<td>0.75</td>
<td>0.73</td>
<td>0.74</td>
</tr>
<tr>
<td>Enumeration</td>
<td>60</td>
<td>0.76</td>
<td>0.93</td>
<td>0.84</td>
</tr>
<tr>
<td>Epanalepsis</td>
<td>60</td>
<td>0.63</td>
<td>0.83</td>
<td>0.72</td>
</tr>
<tr>
<td>Epiphoza</td>
<td>60</td>
<td>0.61</td>
<td>0.93</td>
<td>0.74</td>
</tr>
<tr>
<td>Epizeugma</td>
<td>60</td>
<td>0.68</td>
<td>0.7</td>
<td>0.69</td>
</tr>
<tr>
<td>Epizeuxis</td>
<td>60</td>
<td>0.79</td>
<td>0.77</td>
<td>0.78</td>
</tr>
<tr>
<td>Hypozeugma</td>
<td>60</td>
<td>0.61</td>
<td>0.8</td>
<td>0.69</td>
</tr>
<tr>
<td>If Conditional One</td>
<td>60</td>
<td>0.78</td>
<td>0.78</td>
<td>0.78</td>
</tr>
<tr>
<td>If Conditional Three</td>
<td>60</td>
<td>0.86</td>
<td>0.65</td>
<td>0.74</td>
</tr>
<tr>
<td>If Conditional Two</td>
<td>60</td>
<td>0.82</td>
<td>0.75</td>
<td>0.78</td>
</tr>
<tr>
<td>If Conditional Zero</td>
<td>60</td>
<td>0.71</td>
<td>0.76</td>
<td>0.73</td>
</tr>
<tr>
<td>If Counterfactual</td>
<td>60</td>
<td>0.84</td>
<td>0.87</td>
<td>0.85</td>
</tr>
<tr>
<td>Isocolon</td>
<td>180</td>
<td>0.57</td>
<td>0.83</td>
<td>0.68</td>
</tr>
<tr>
<td>Mesarchia</td>
<td>20</td>
<td>0.45</td>
<td>0.85</td>
<td>0.59</td>
</tr>
<tr>
<td>Mesodiplosis</td>
<td>40</td>
<td>0.28</td>
<td>0.68</td>
<td>0.4</td>
</tr>
<tr>
<td>Passive Voice</td>
<td>60</td>
<td>0.79</td>
<td>0.98</td>
<td>0.87</td>
</tr>
<tr>
<td>Polysyndeton</td>
<td>60</td>
<td>0.77</td>
<td>0.7</td>
<td>0.73</td>
</tr>
<tr>
<td>Pysma</td>
<td>60</td>
<td>1</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>Superlative Adjective</td>
<td>70</td>
<td>0.62</td>
<td>0.73</td>
<td>0.67</td>
</tr>
<tr>
<td>Superlative Adverb</td>
<td>70</td>
<td>0.63</td>
<td>0.5</td>
<td>0.56</td>
</tr>
<tr>
<td>Unless Conditional</td>
<td>60</td>
<td>1</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>Whether Conditional</td>
<td>60</td>
<td>1</td>
<td>0.83</td>
<td>0.91</td>
</tr>
</tbody>
</table>

- Balance schemes
- Omission schemes
- Repetition schemes
- Custom schemes
If-conditional Detection

If P, then Q

Stanford Dependencies

Governors

If ... governor

P-clause

If-conditional

... governor

Q-clause

grammatical rules

4 tokens
If-counterfactual Detection

If P, then Q

Identify P and Q

P → past tense (VBN/VBD)

Q → modals (would/could…)

If-counterfactual
Presidential Debates: Findings

<table>
<thead>
<tr>
<th>Debate Type</th>
<th>Distribution (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Clinton → Trump</td>
<td>11.00</td>
</tr>
<tr>
<td>Trump → Clinton</td>
<td>7.02</td>
</tr>
</tbody>
</table>