Mining Rhetorical Devices by means of Natural Language Processing

Bauhaus-Universität Weimar

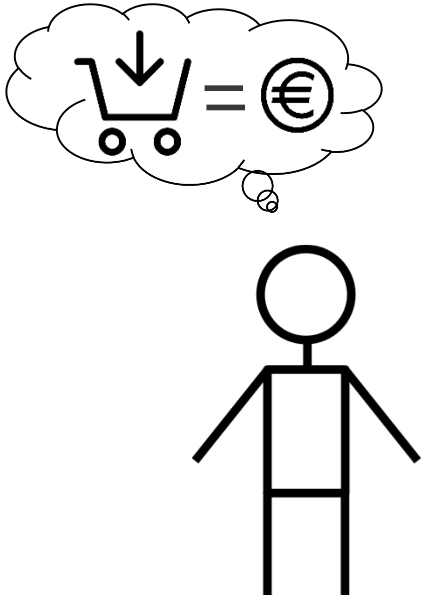
Viorel Morari
viorel.morari@uni-weimar.de

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

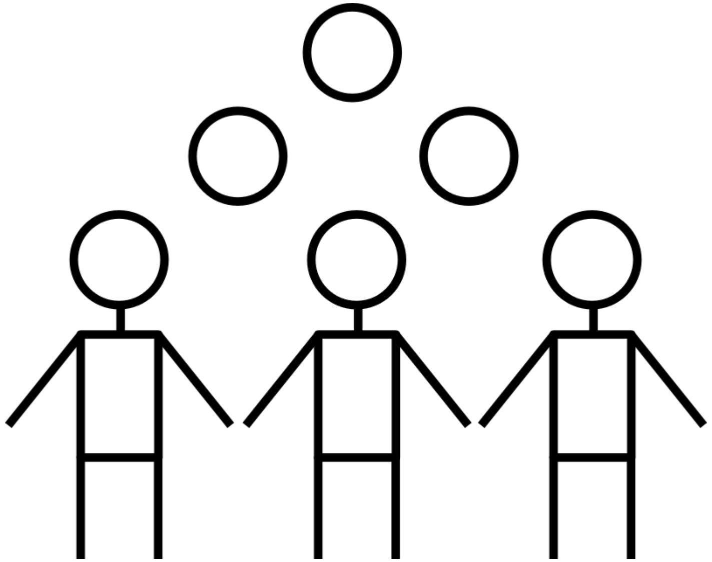
Master Thesis Defense
January 23rd , 2018

Advisor: Khalid Al-Khatib

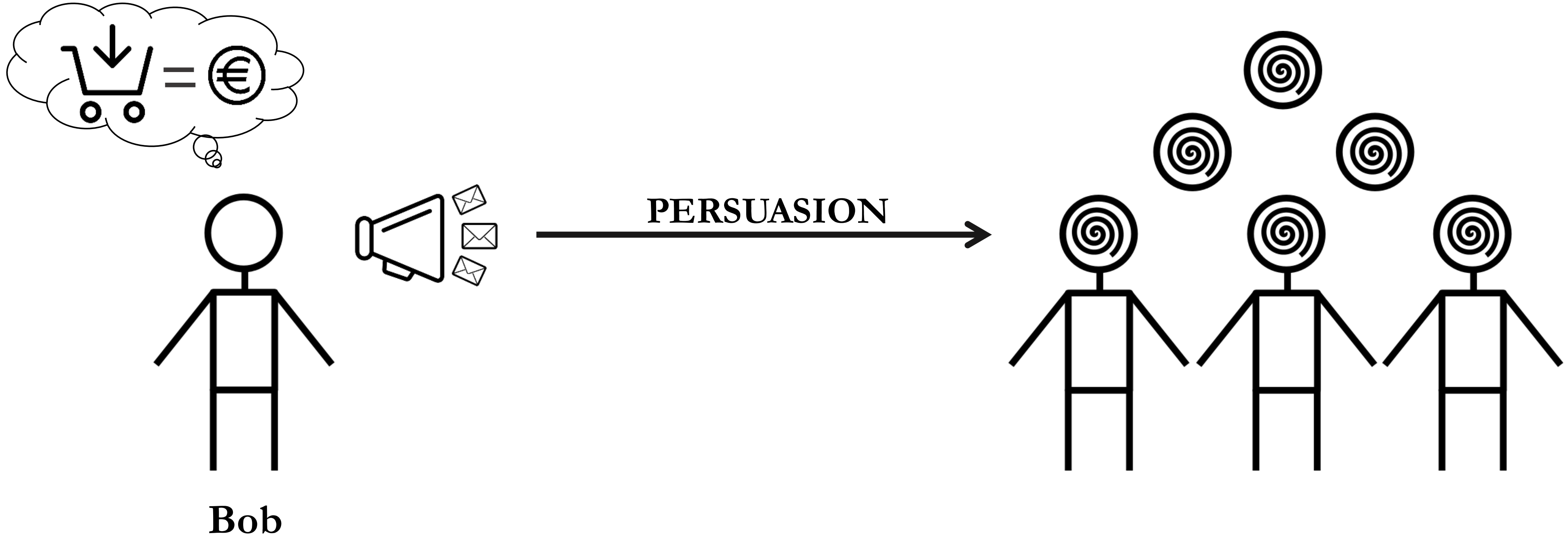
What is Rhetoric?



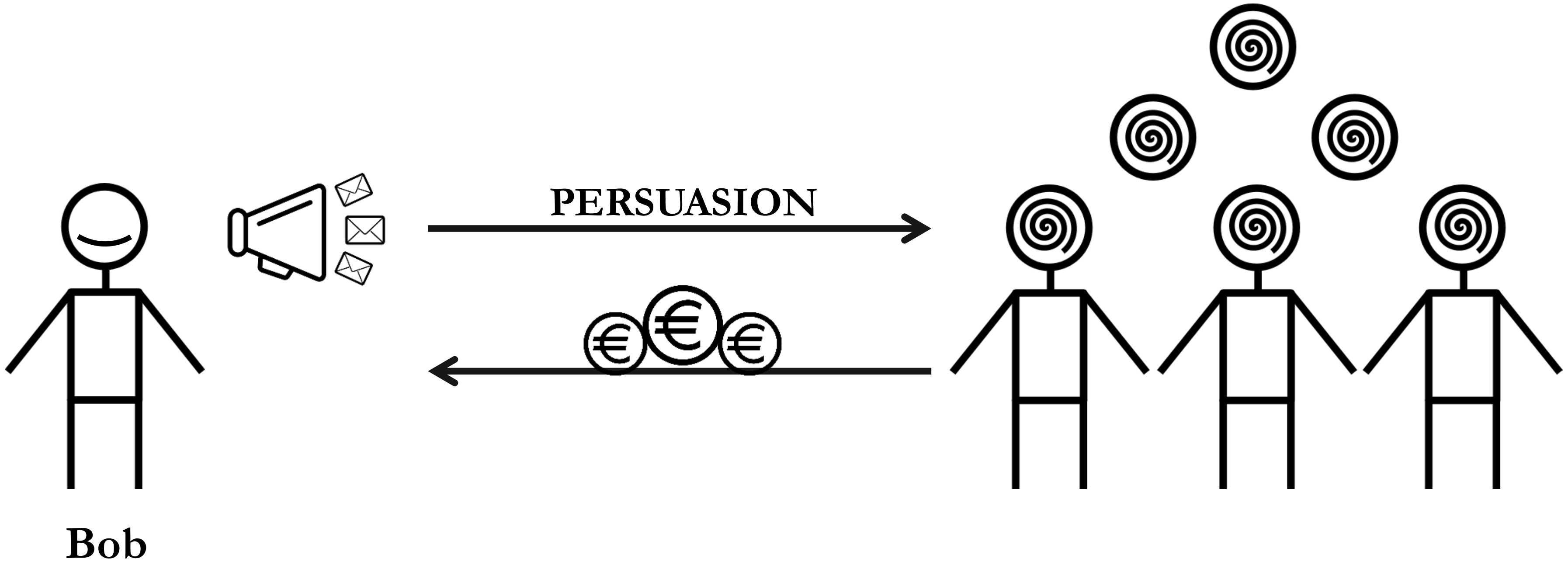
Bob



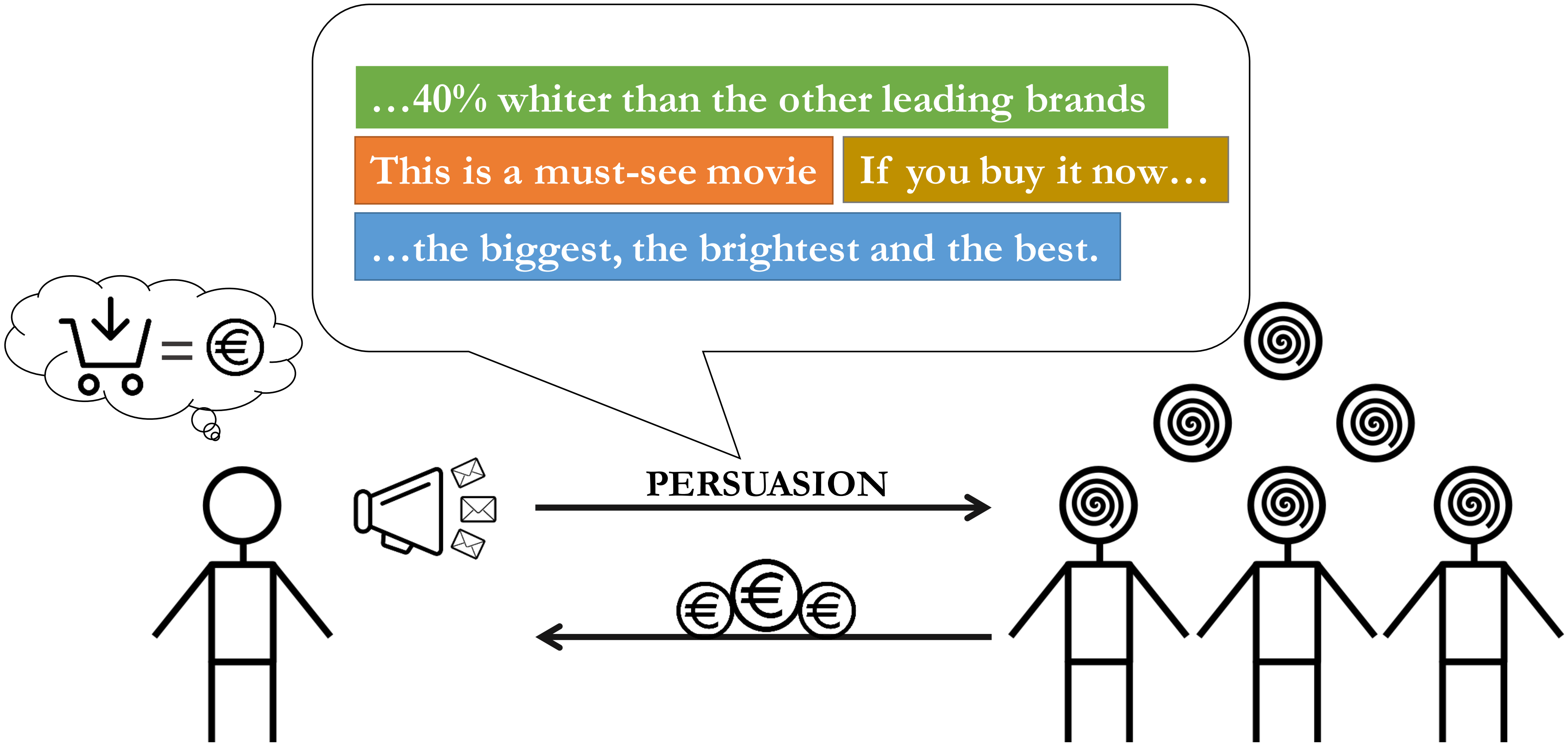
What is Rhetoric?



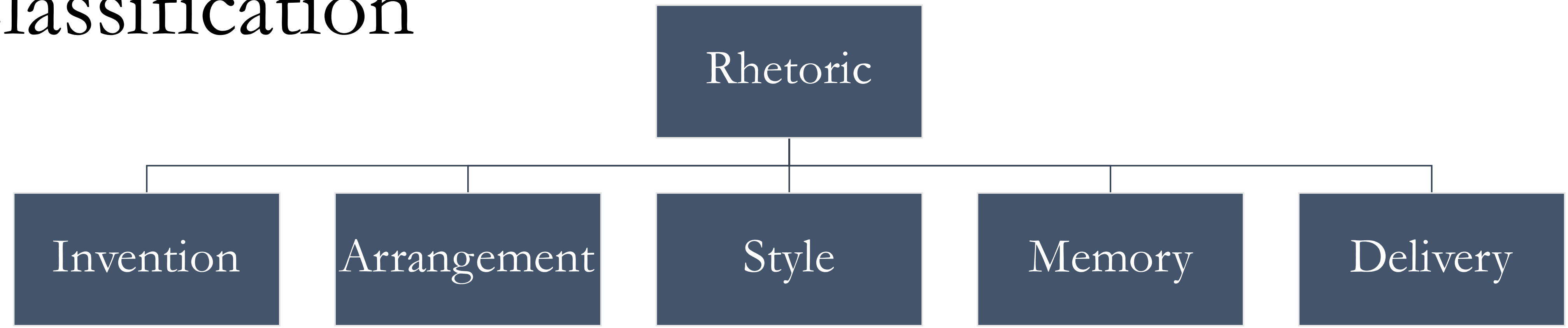
What is Rhetoric?



What is a Rhetorical Device?

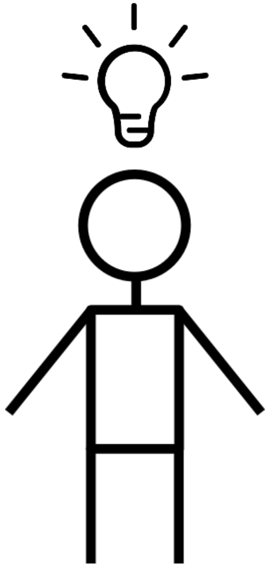


Classification



Classification

- Rhetoric
  - Invention
  - Arrangement
  - Style
  - Memory
  - Delivery



Bob

Happiness

Feelings

Sad

Drink

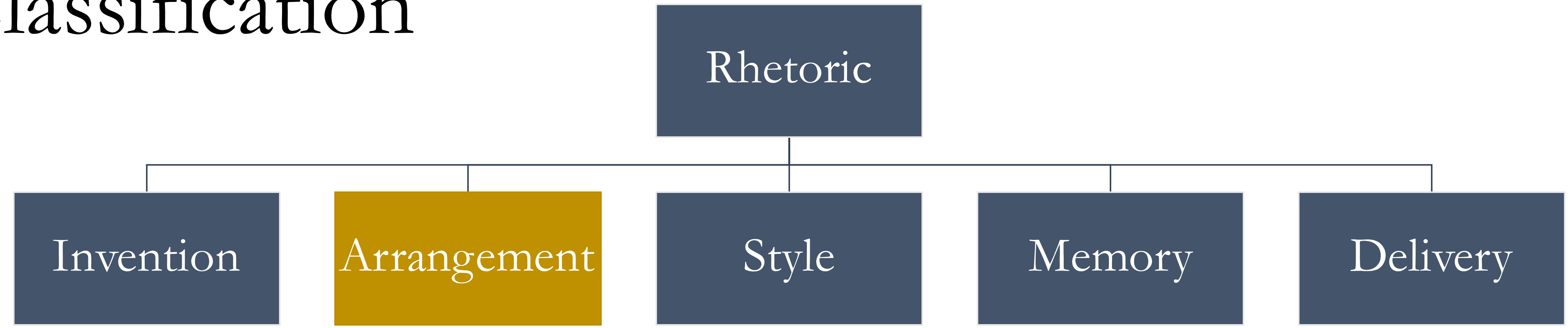
Bottle

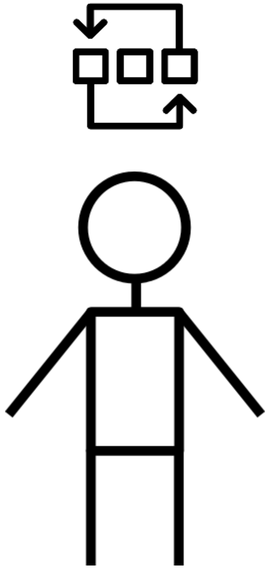
Open happiness.

Grab a bottle.

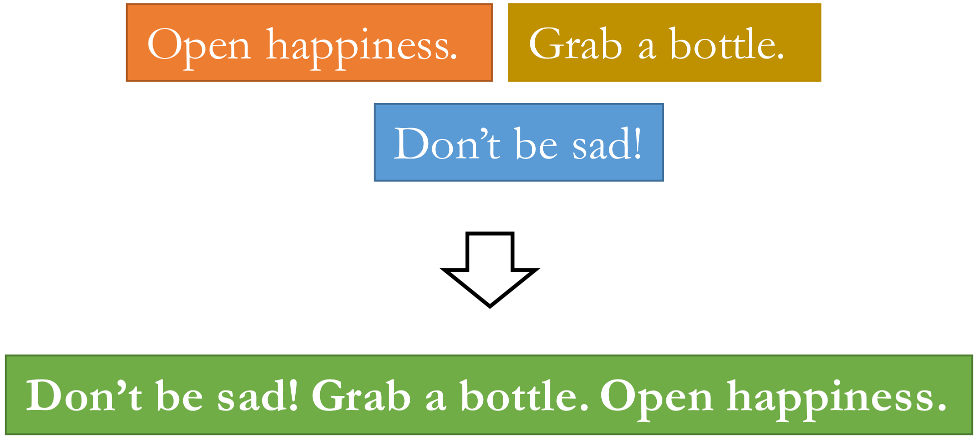
Don't be sad!

Classification





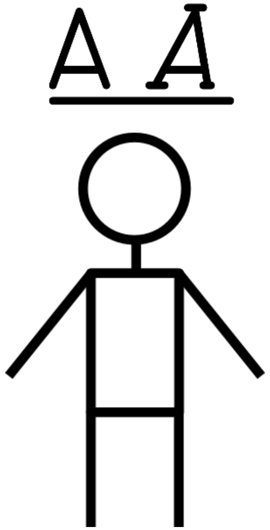
Bob



Classification

Rhetoric

- Invention
- Arrangement
- Style
- Memory
- Delivery



Bob

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

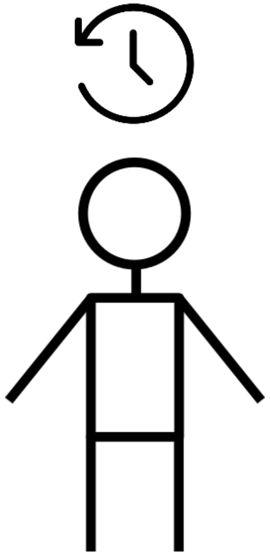
Feeling down? Open a bottle, open happiness!

Rhetorical question

Repetition, Balance

Classification

- Rhetoric
  - Invention
  - Arrangement
  - Style
  - **Memory**
  - Delivery



Bob

Feeling down? Open a bottle, open happiness!

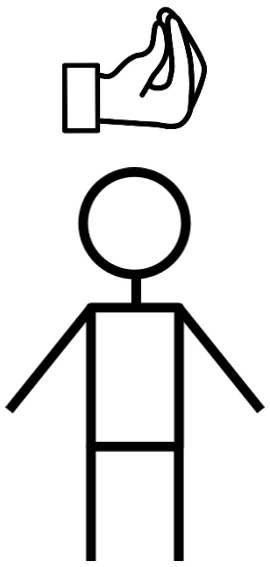
Feeling down? Open a bottle, open happiness!

Feeling down? Open a bottle, open happiness!

Classification

Rhetoric

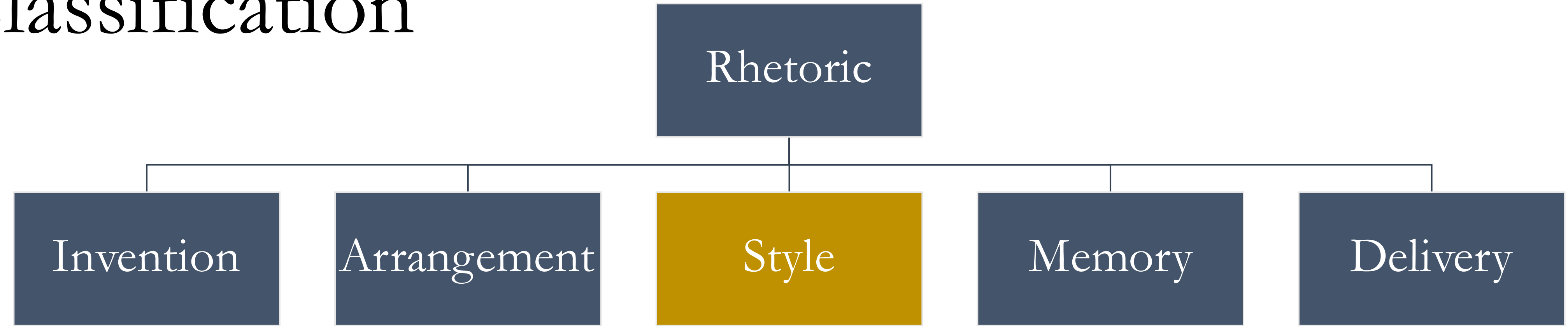
- Invention
- Arrangement
- Style
- Memory
- Delivery



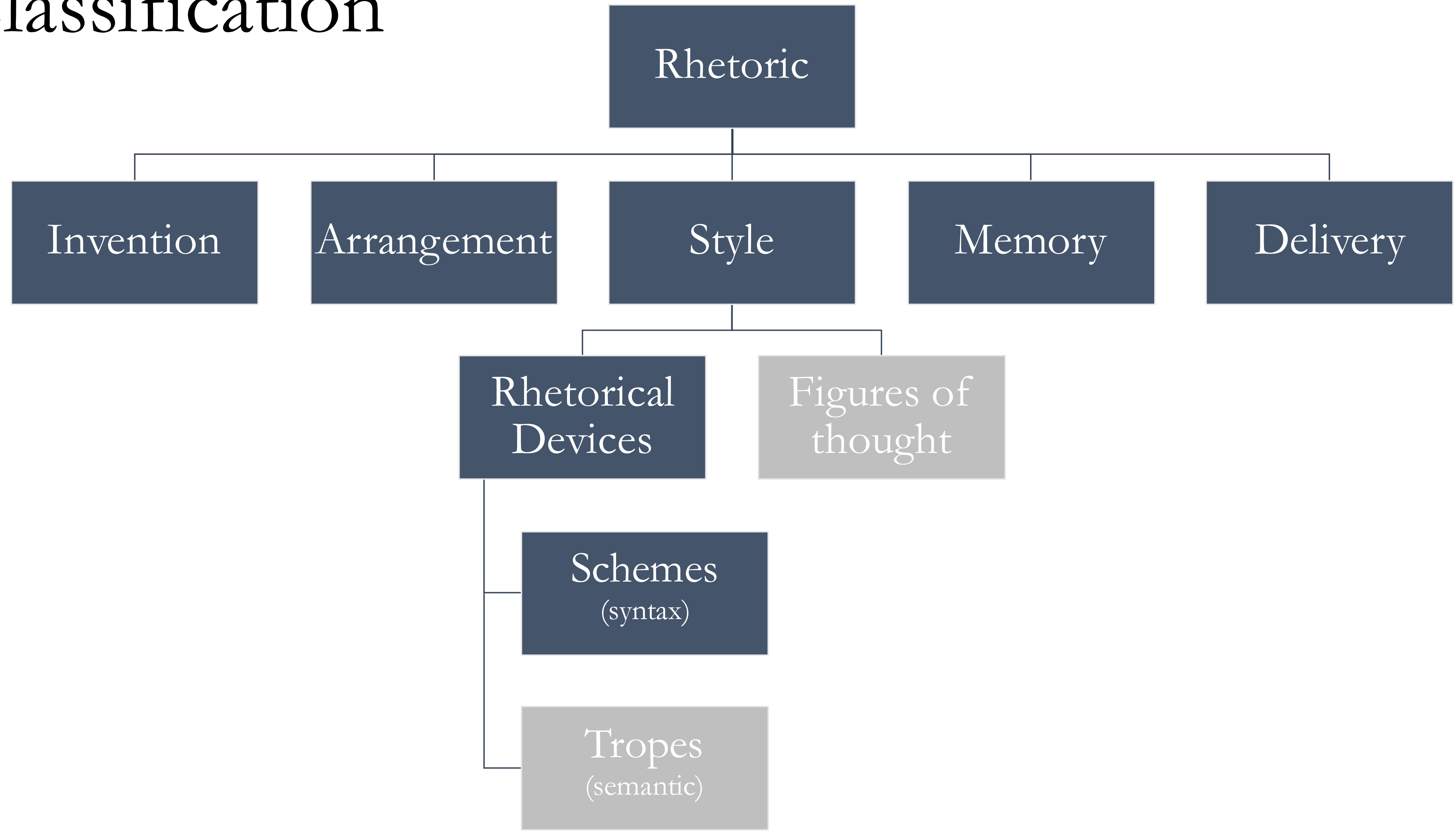
Bob

Feeling down? Open a bottle, open happiness!

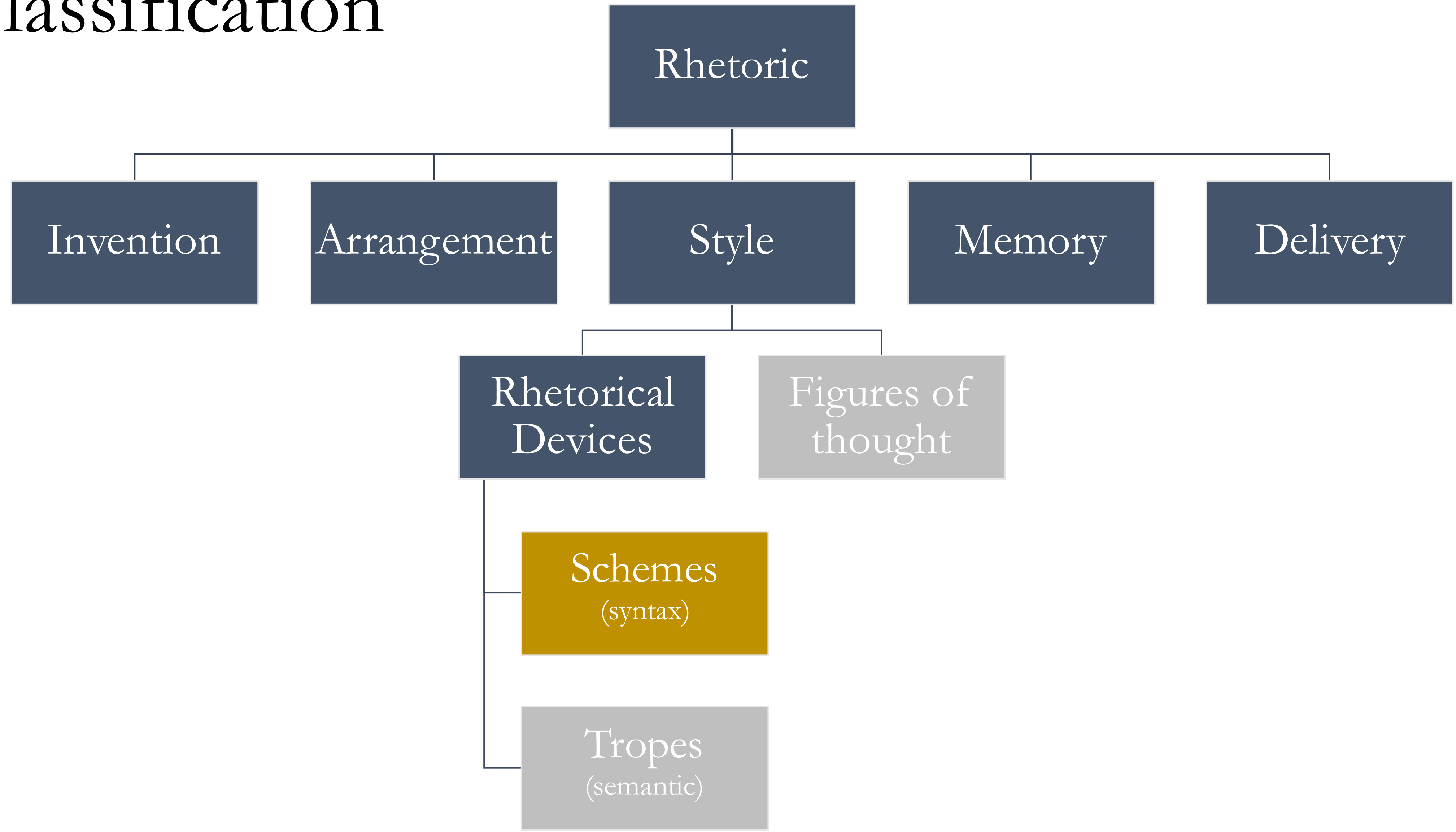
Classification



Classification

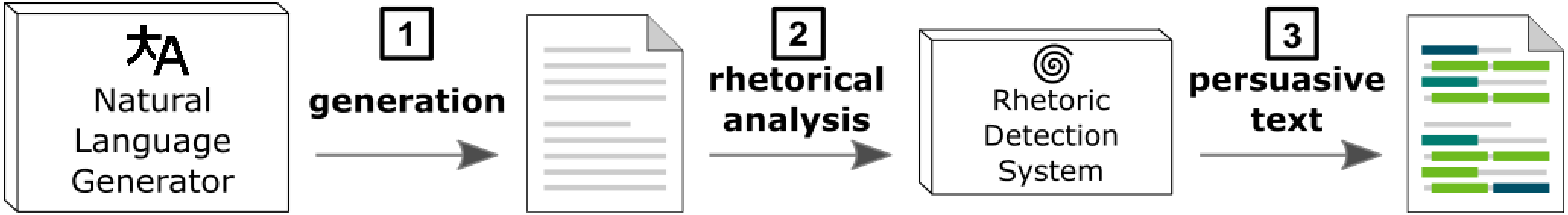


Classification

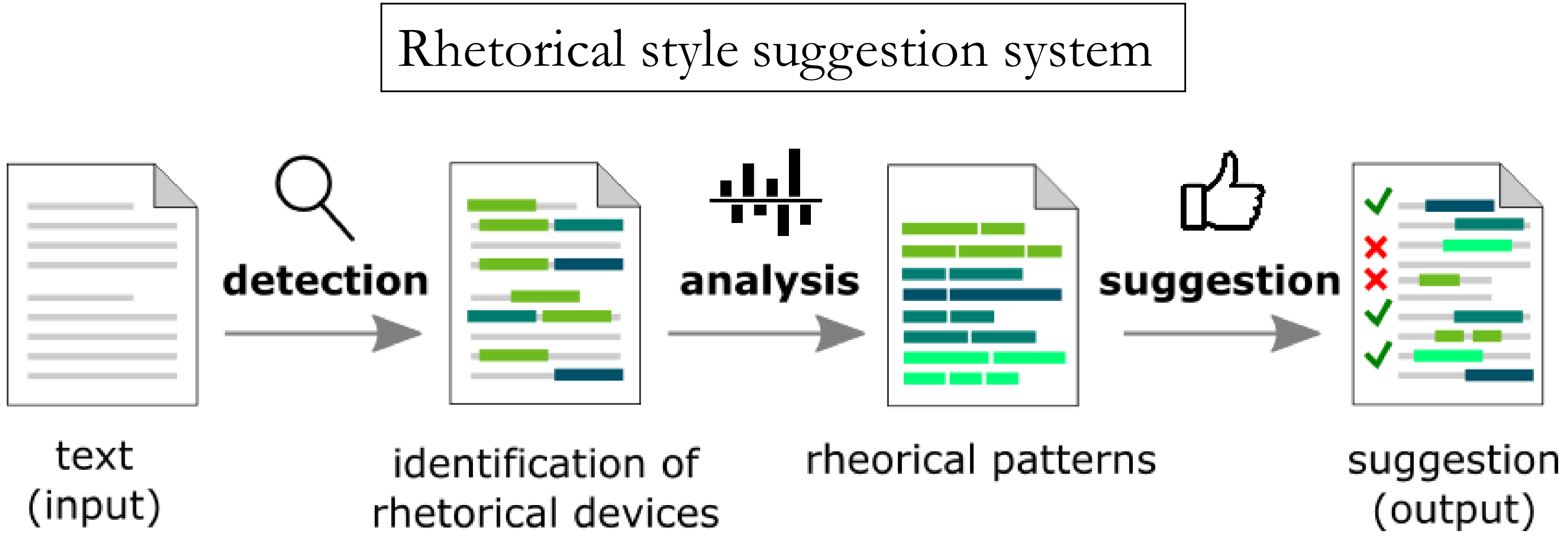


Envisioned Applications

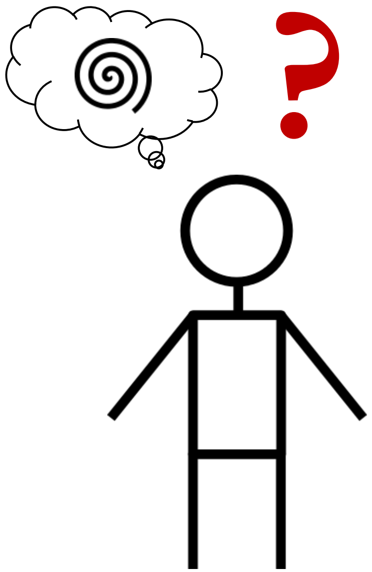
Rhetoric-based NLG system



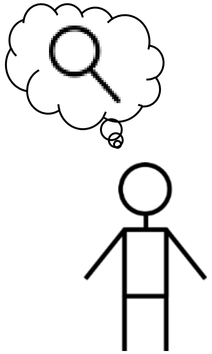
Envisioned Applications

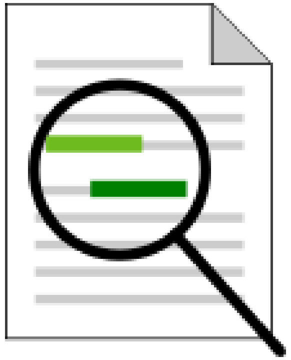


Research Questions



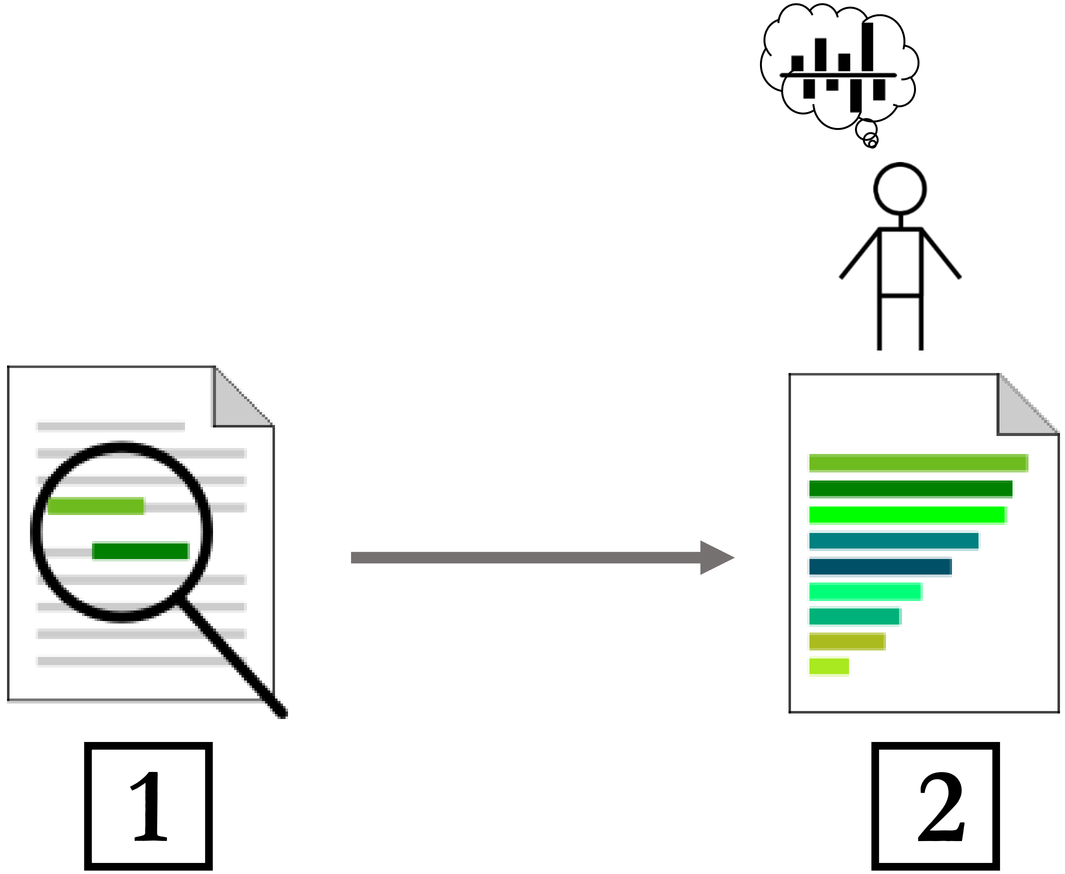
Research Questions



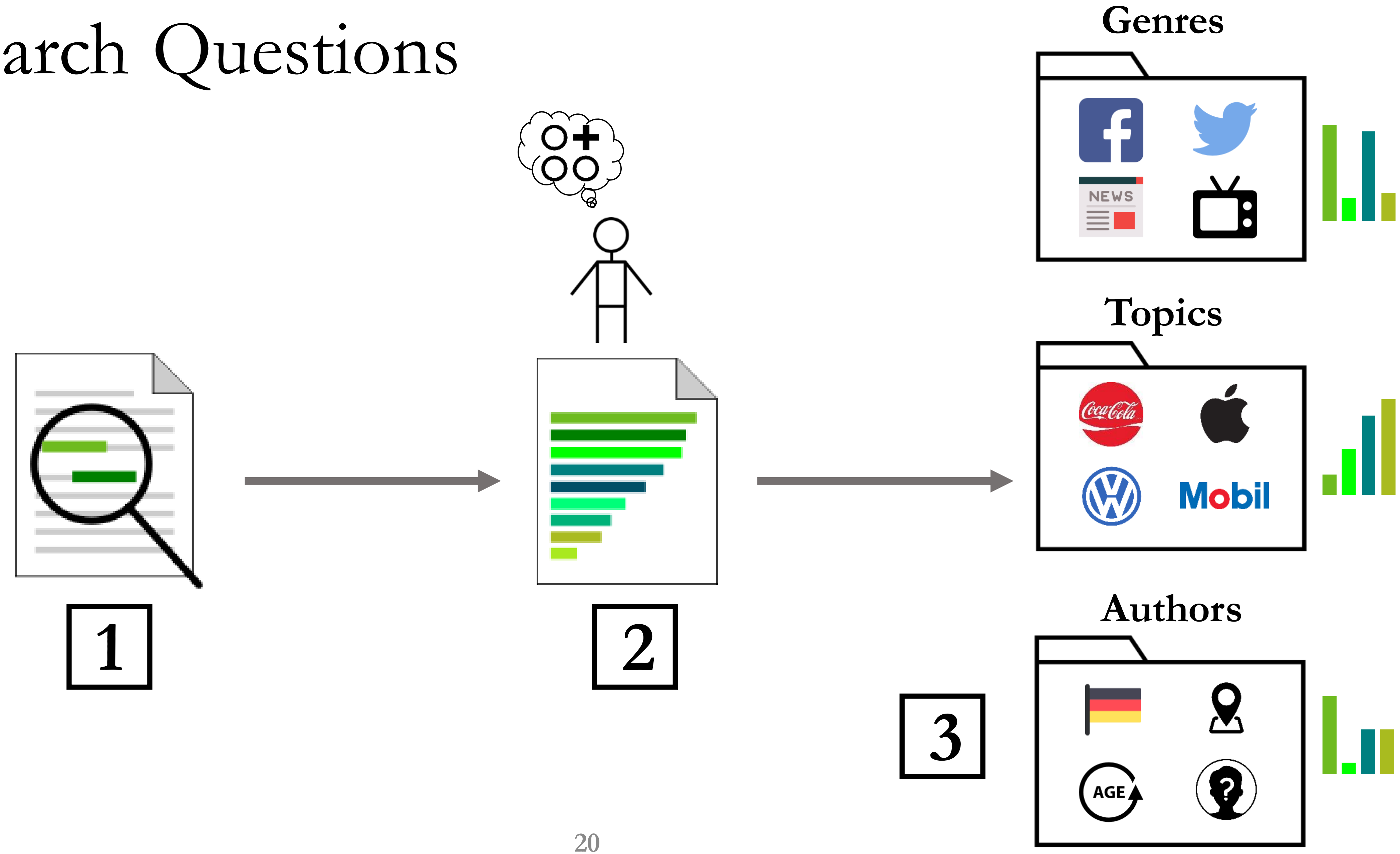


1

Research Questions

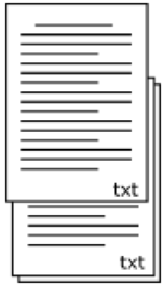


Research Questions



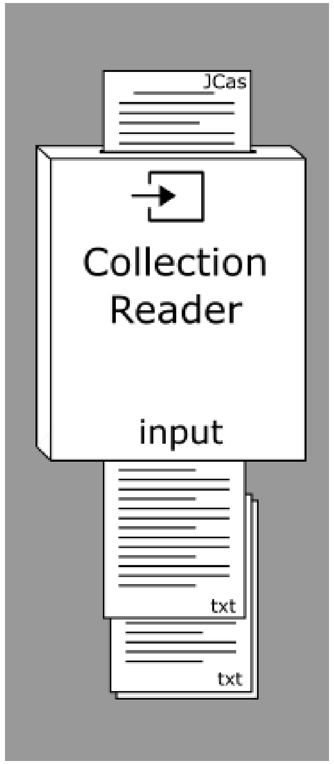
1 Detection of Rhetorical Devices

Pipeline – UIMA

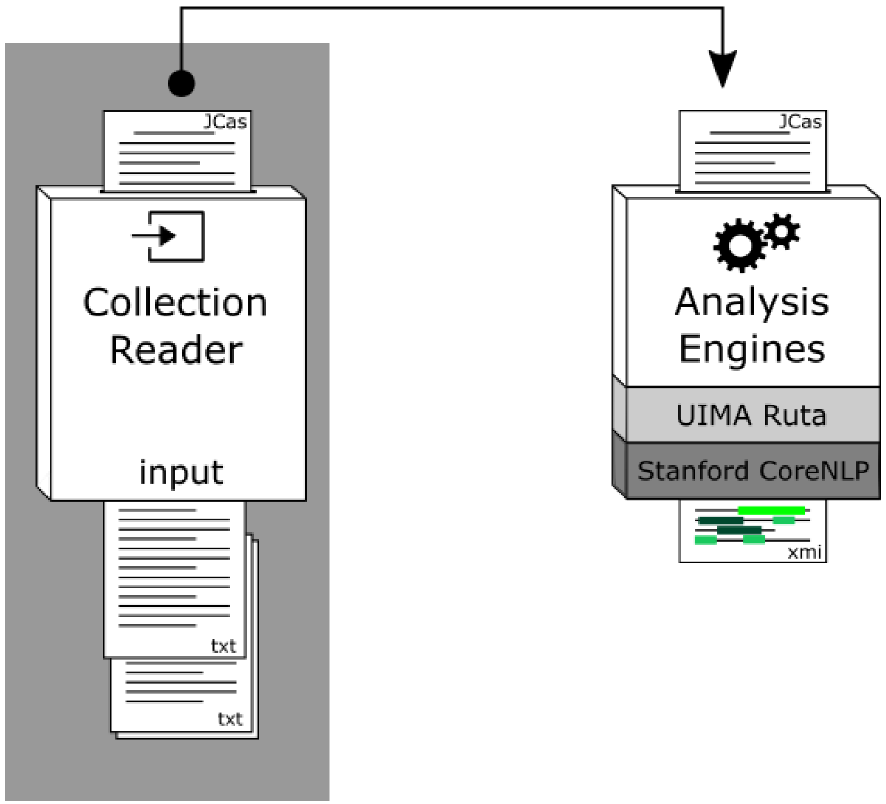


input

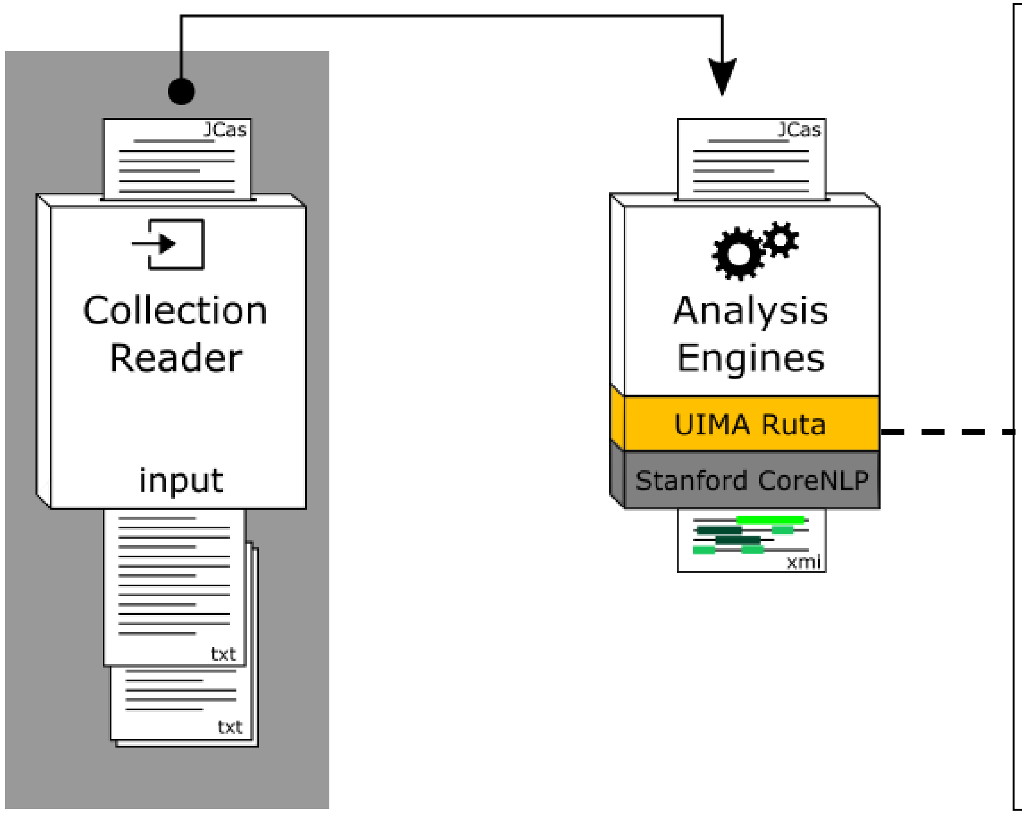
Pipeline – UIMA



Pipeline – UIMA



Pipeline – UIMA Ruta



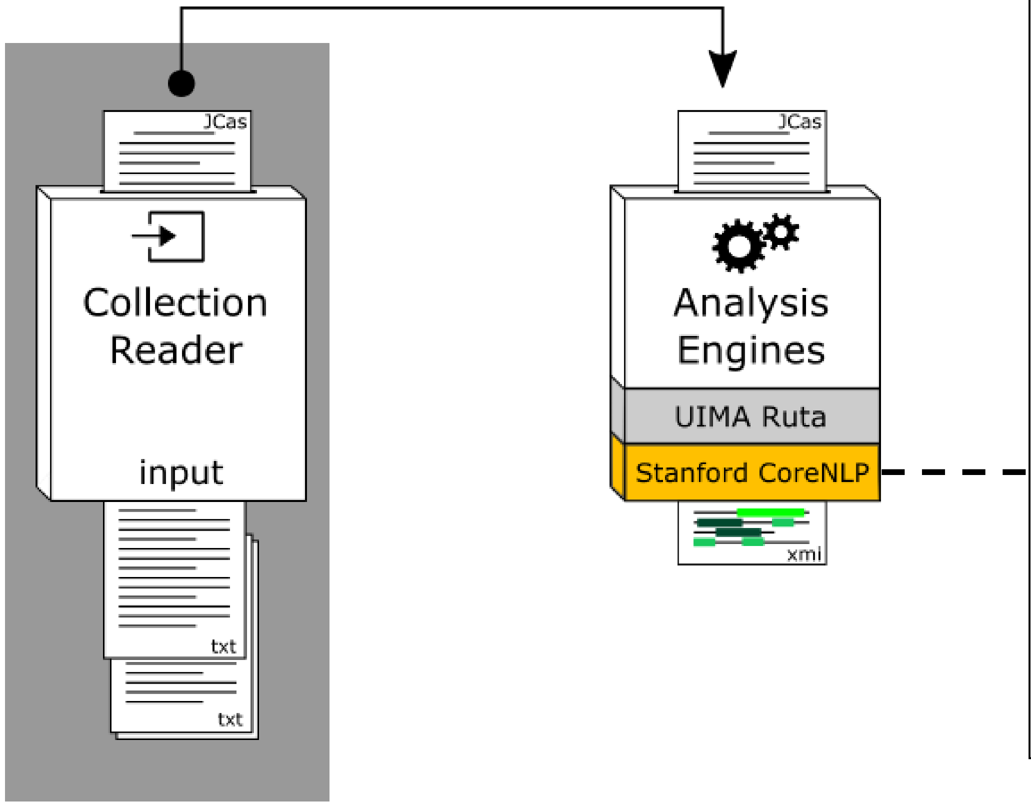
- **UIMA Rule-based Text Annotation** - intuitive and flexible domain specific language for defining patterns of annotations (Kluegl et al. [2016]).

- Example:

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

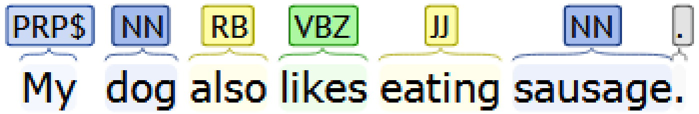
… . This is a sample sentence. …

Pipeline – Stanford CoreNLP

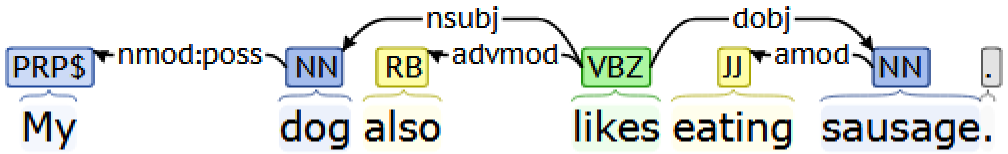


- **Stanford CoreNLP** – a suite of tools for linguistic analysis.

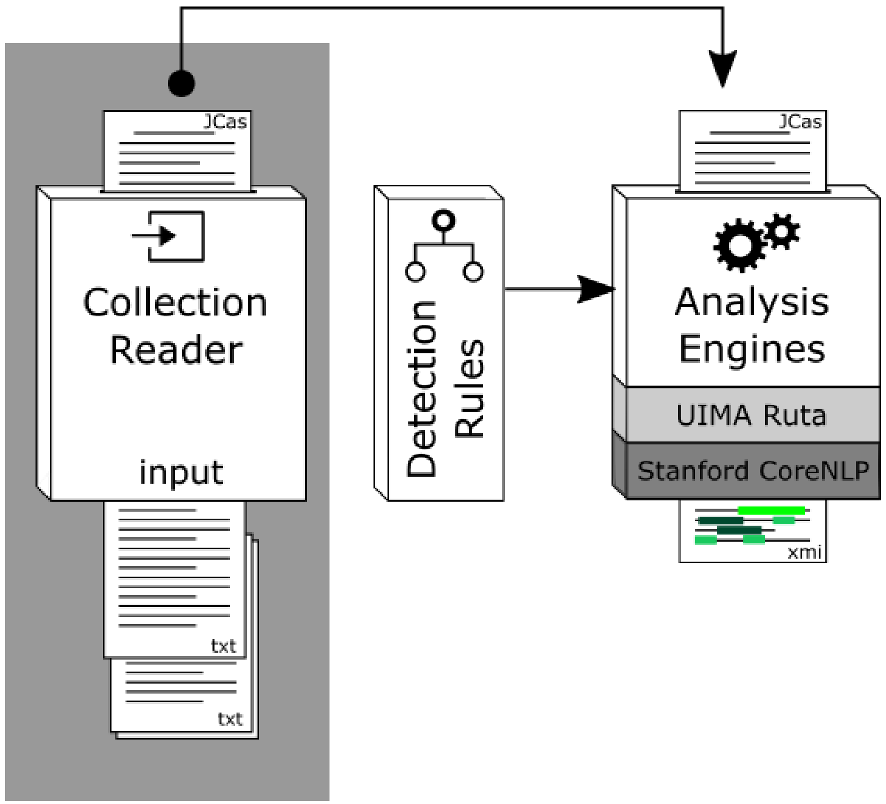
- We use:
  - Stanford Parser



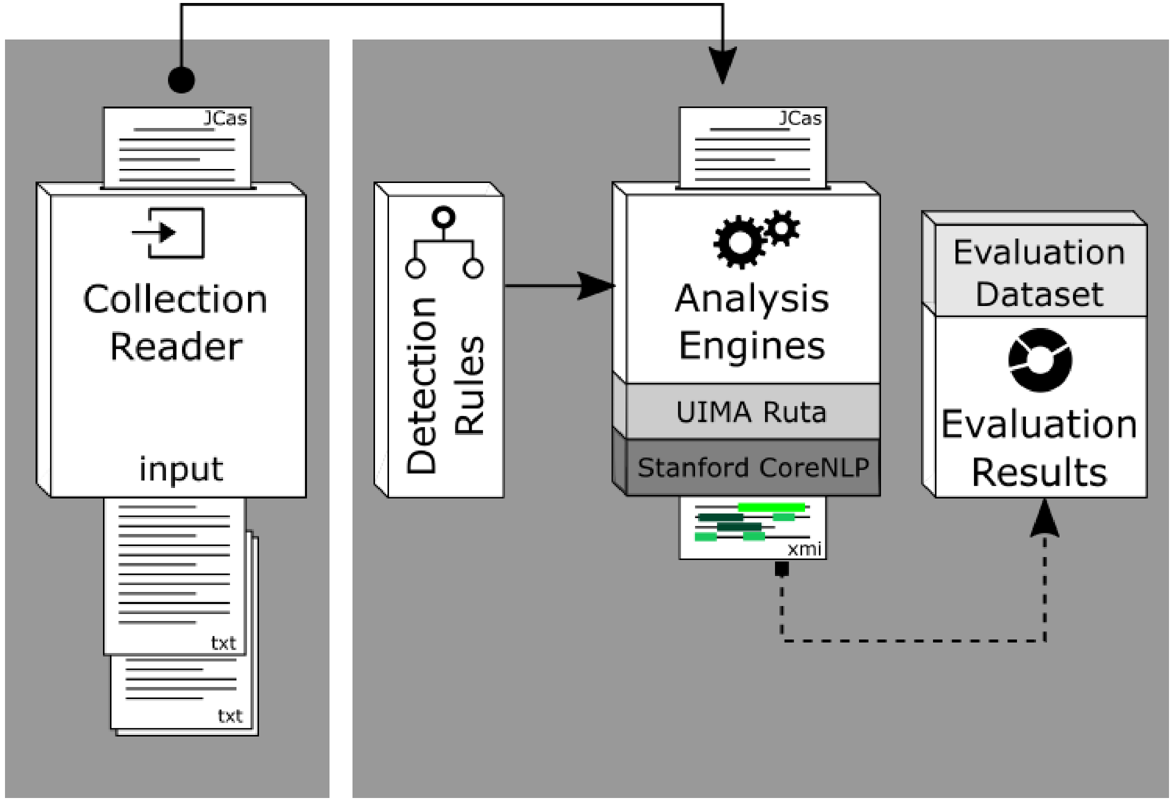
  - Stanford Dependencies



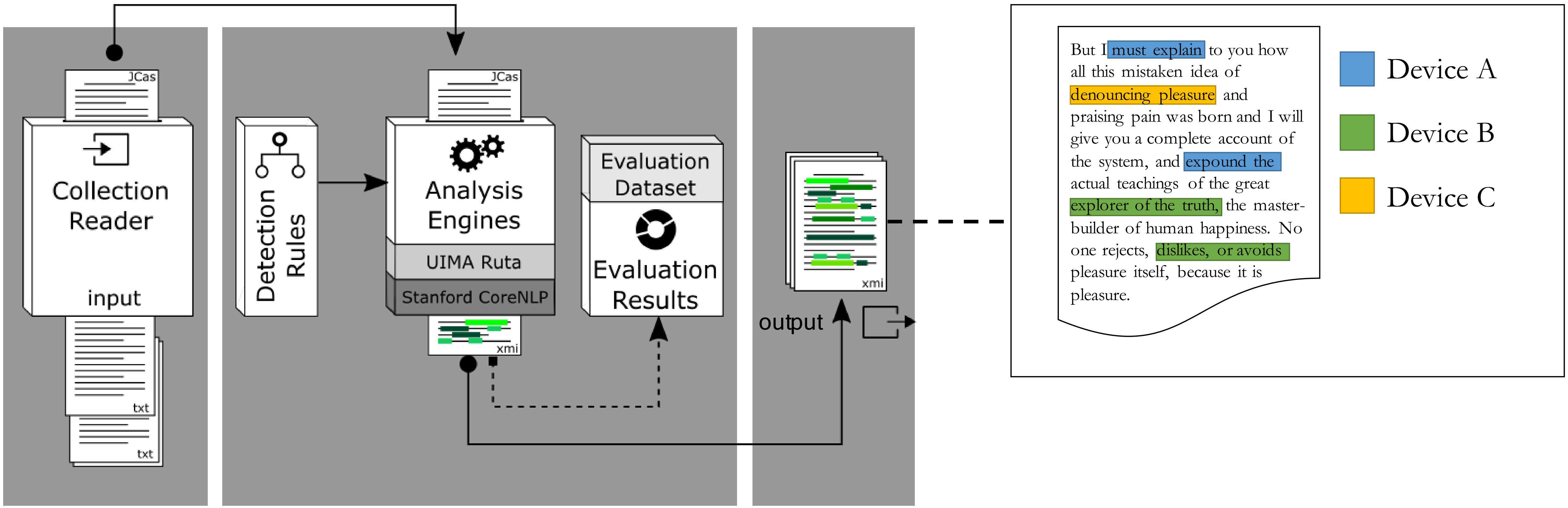
Pipeline – UIMA



Pipeline – UIMA



Pipeline – UIMA



Rhetorical Devices

Balance schemes

Omission schemes

Repetition schemes

Custom schemes

Interplay between equivalent ideas

Control the rhythm of thought

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	Omission schemes	Repetition schemes	Custom schemes
• Enumeration • Pysma • Isocolon -bicolon -tricolon -tetracolon	• Asyndeton • Hypozeugma • Epizeugma	• Epanalepsis • Mesarchia • Epiphoza • Mesodiplosis • Anadiplosis • Diacope • Epizeuxis • Polysyndeton	• 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)

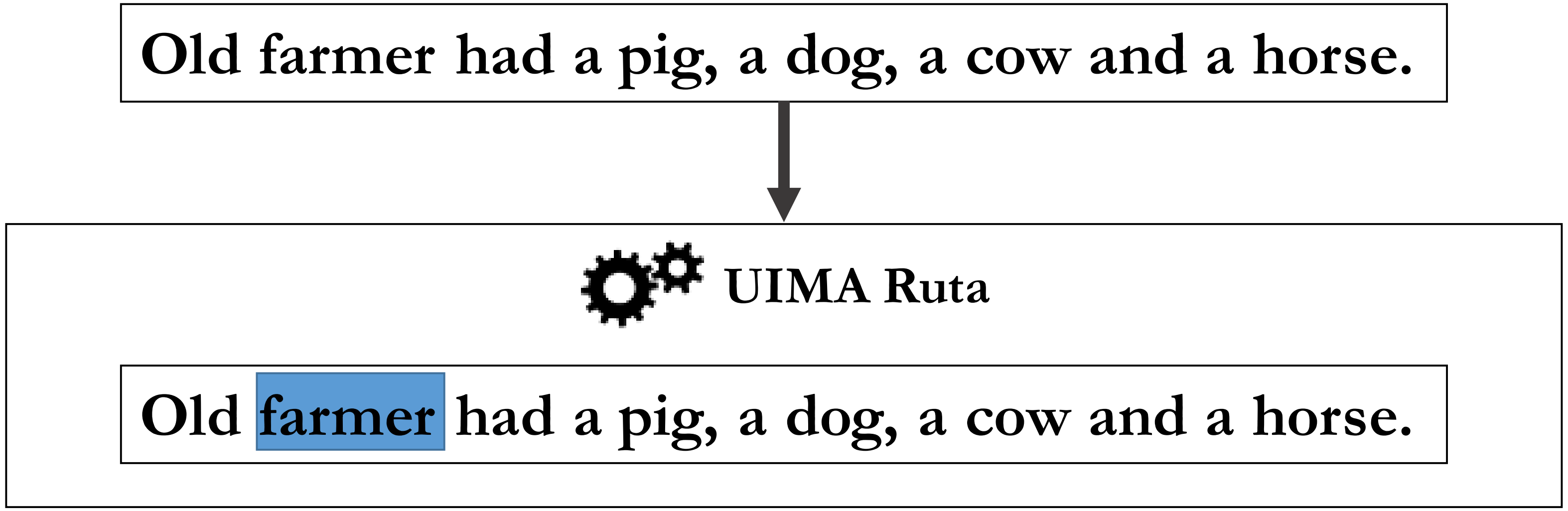
Old farmer had a pig, a dog, a cow and a horse.

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)



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)

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)

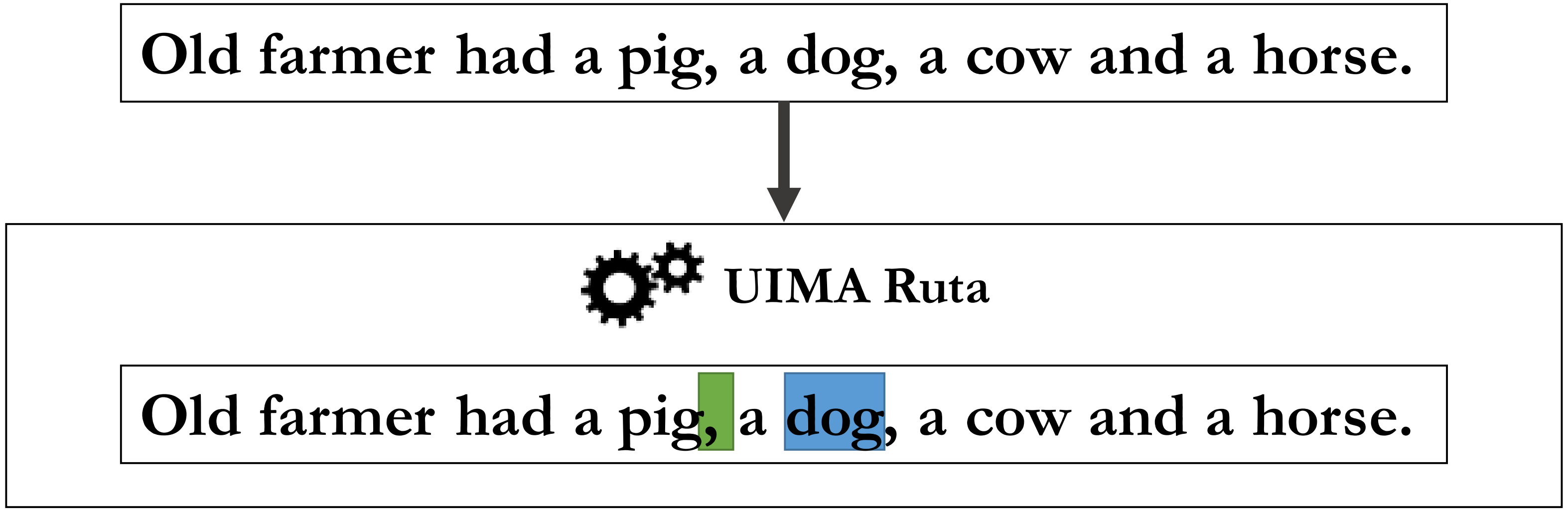
Old farmer had a pig, a dog, a cow and a horse.

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)



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)

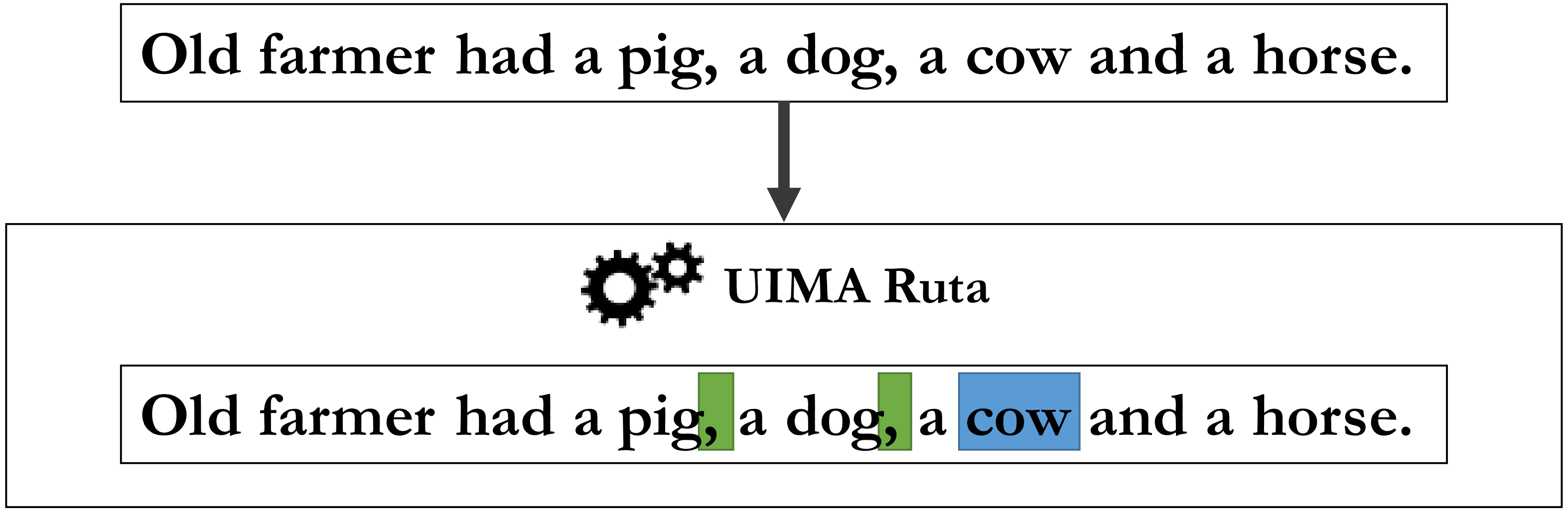
Old farmer had a pig, a dog, a cow and a horse.

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)



Balance: Enumeration

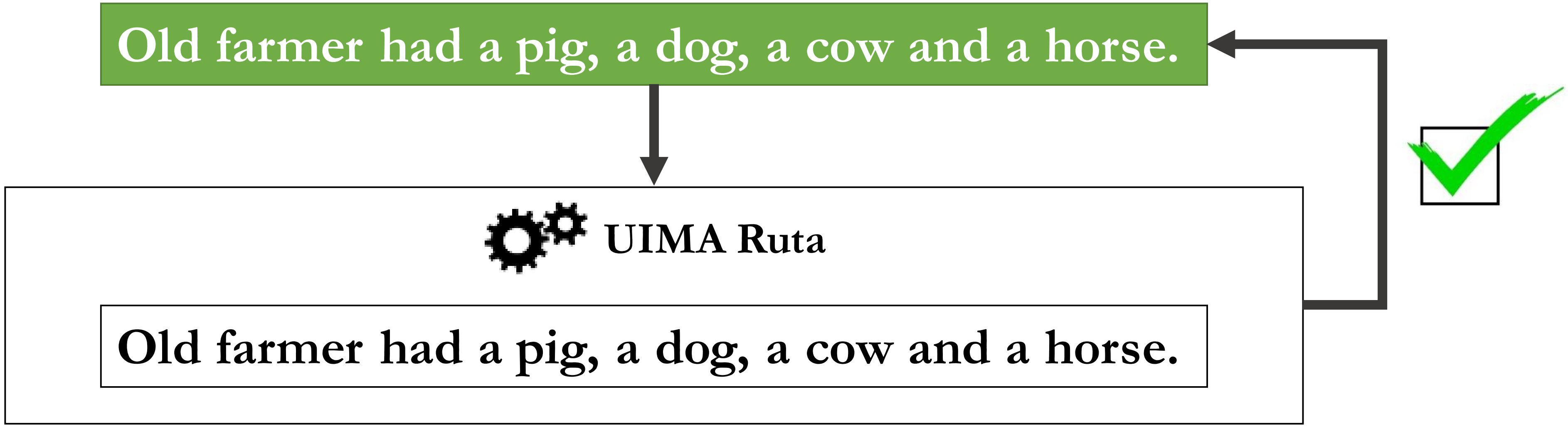
Old farmer had a pig, a dog, a cow and a horse.

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)



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)

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)

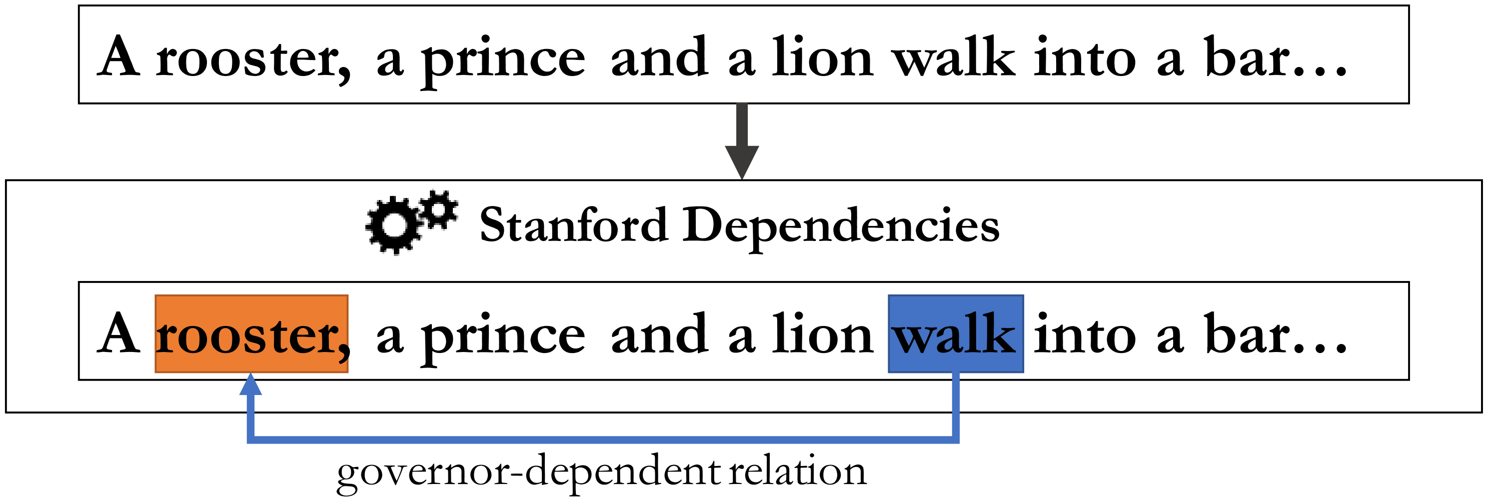
A rooster, a prince and a lion walk into a bar…

Stanford Dependencies

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)

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

Stanford Dependencies

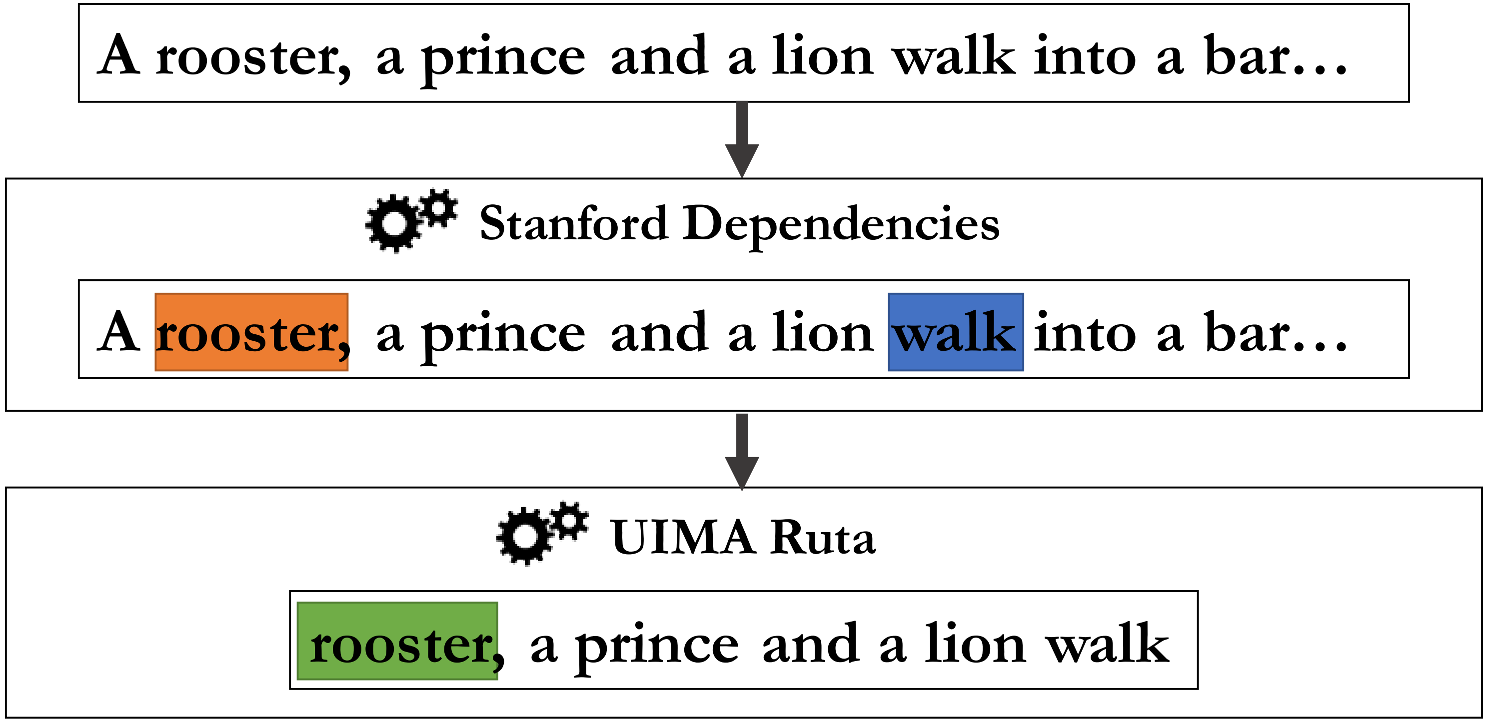
A rooster, a prince and a lion walk into a bar…

UIMA Ruta

rooster, a prince and a lion walk

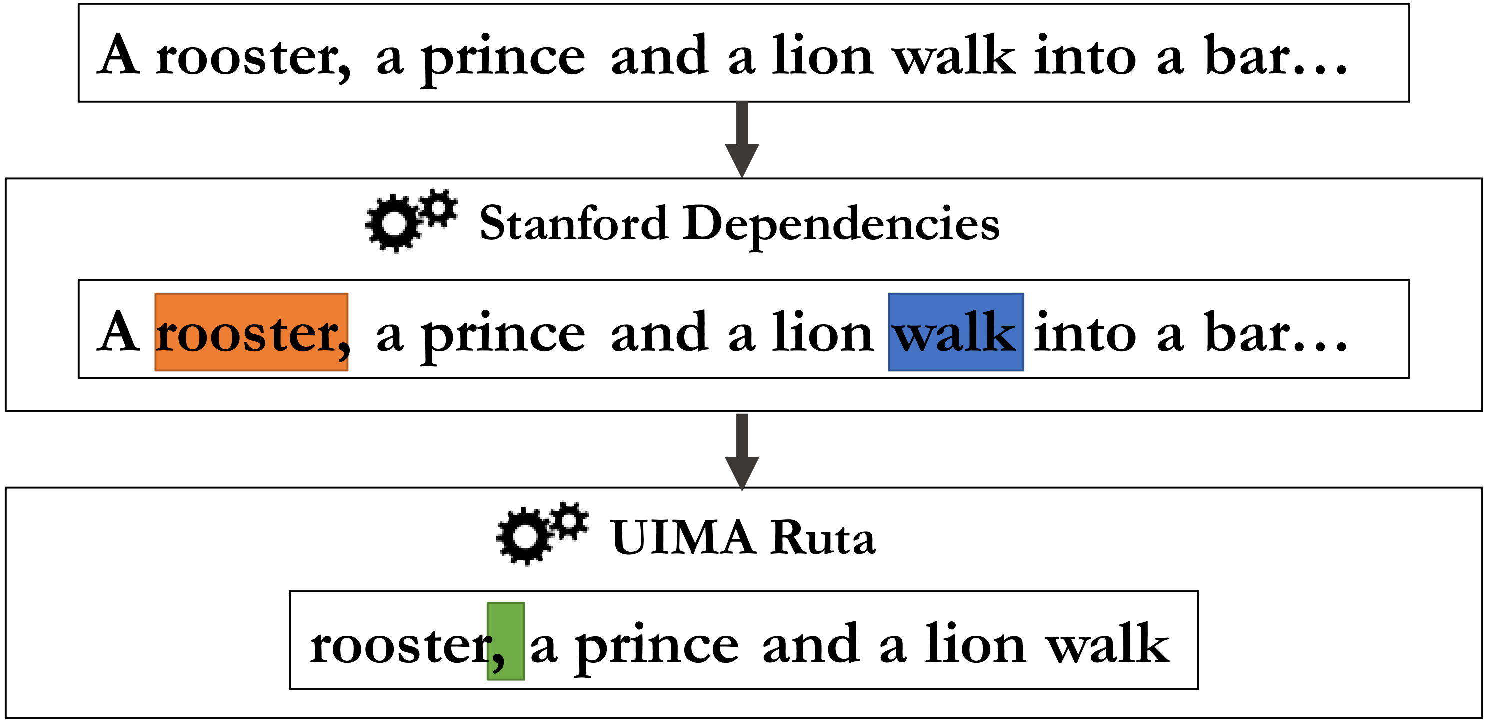
Omission: Hypozeugma

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)



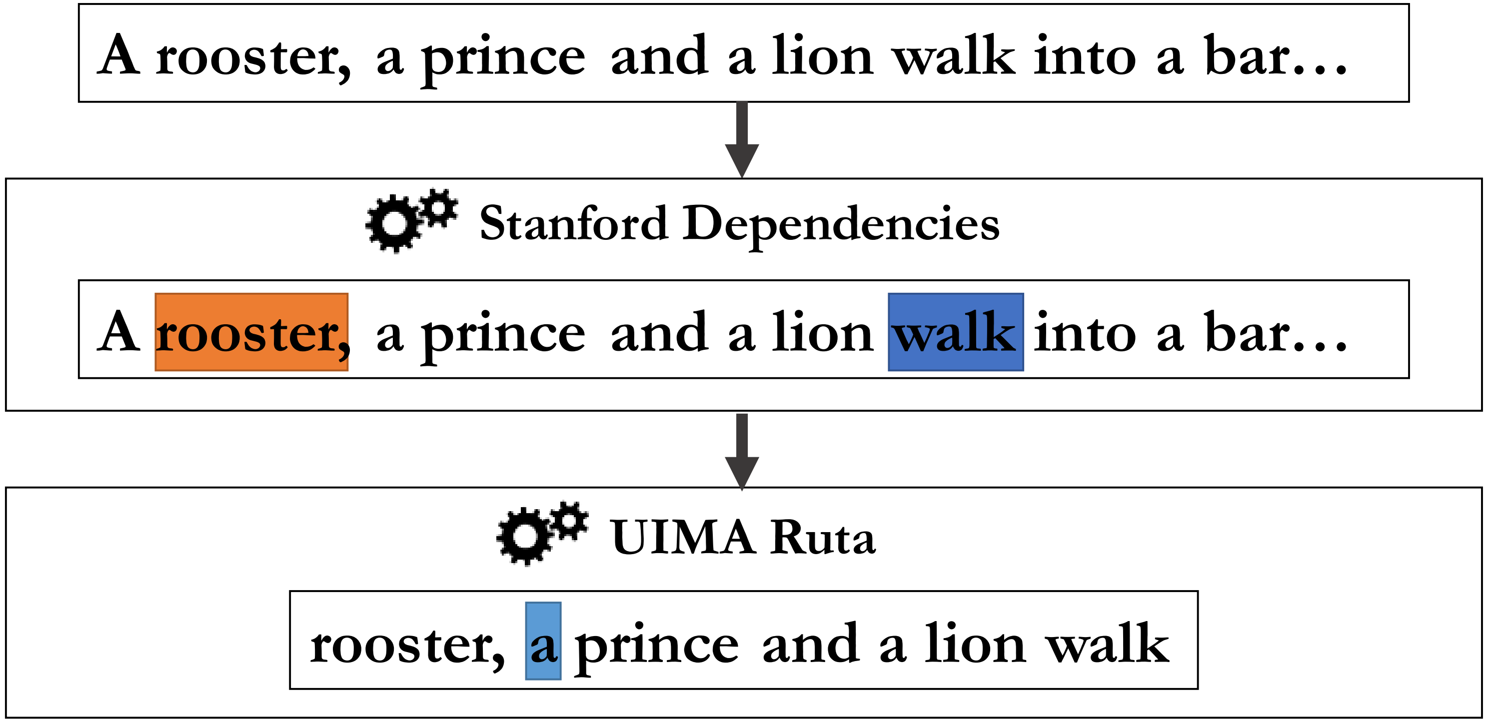
Omission: Hypozeugma

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)



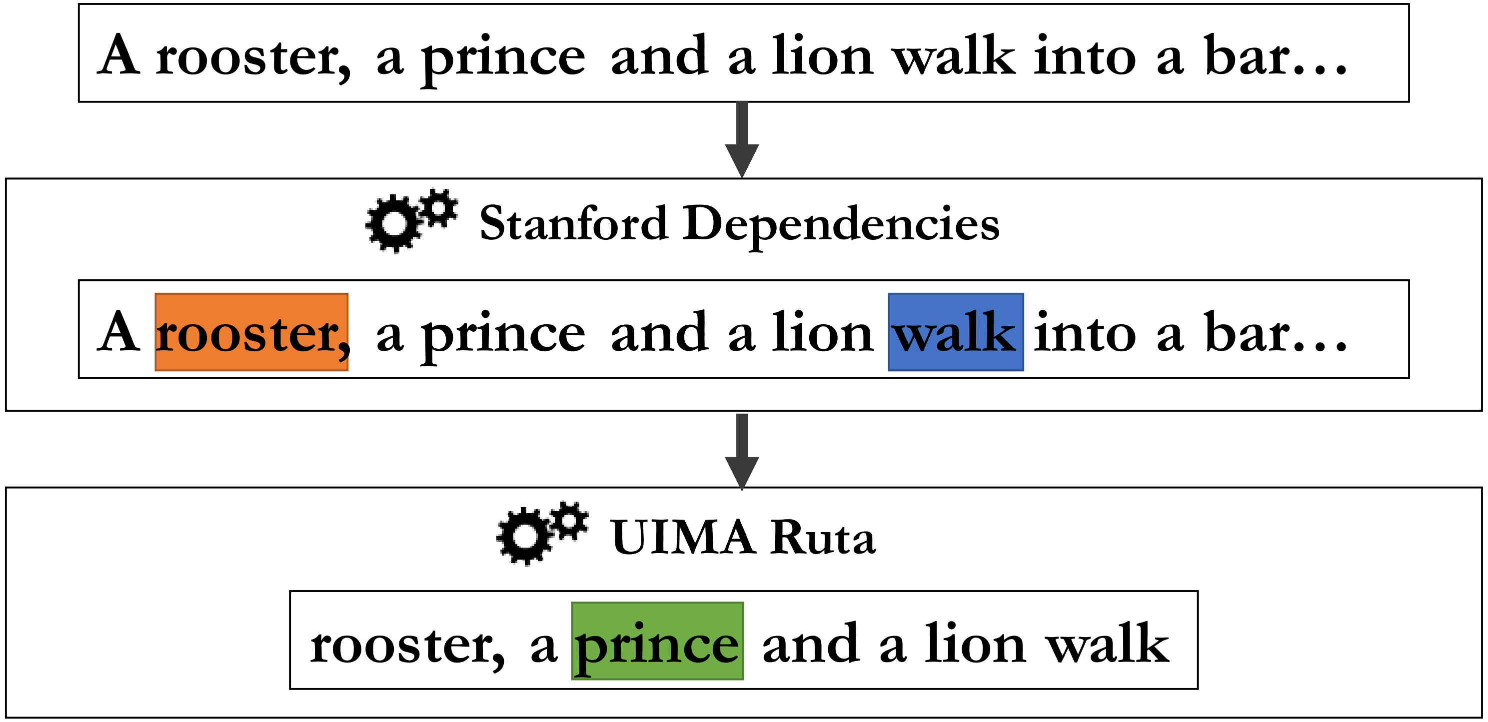
Omission: Hypozeugma

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)



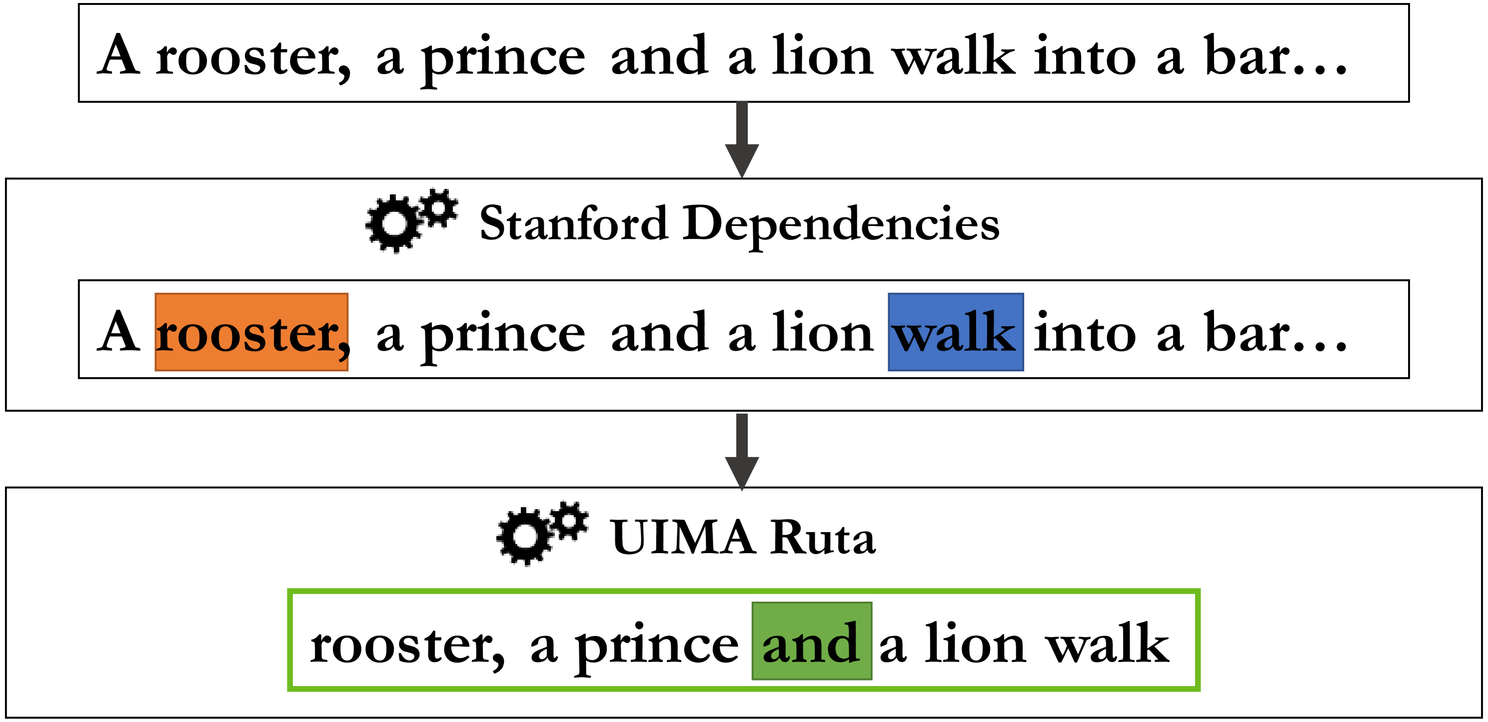
Omission: Hypozeugma

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)



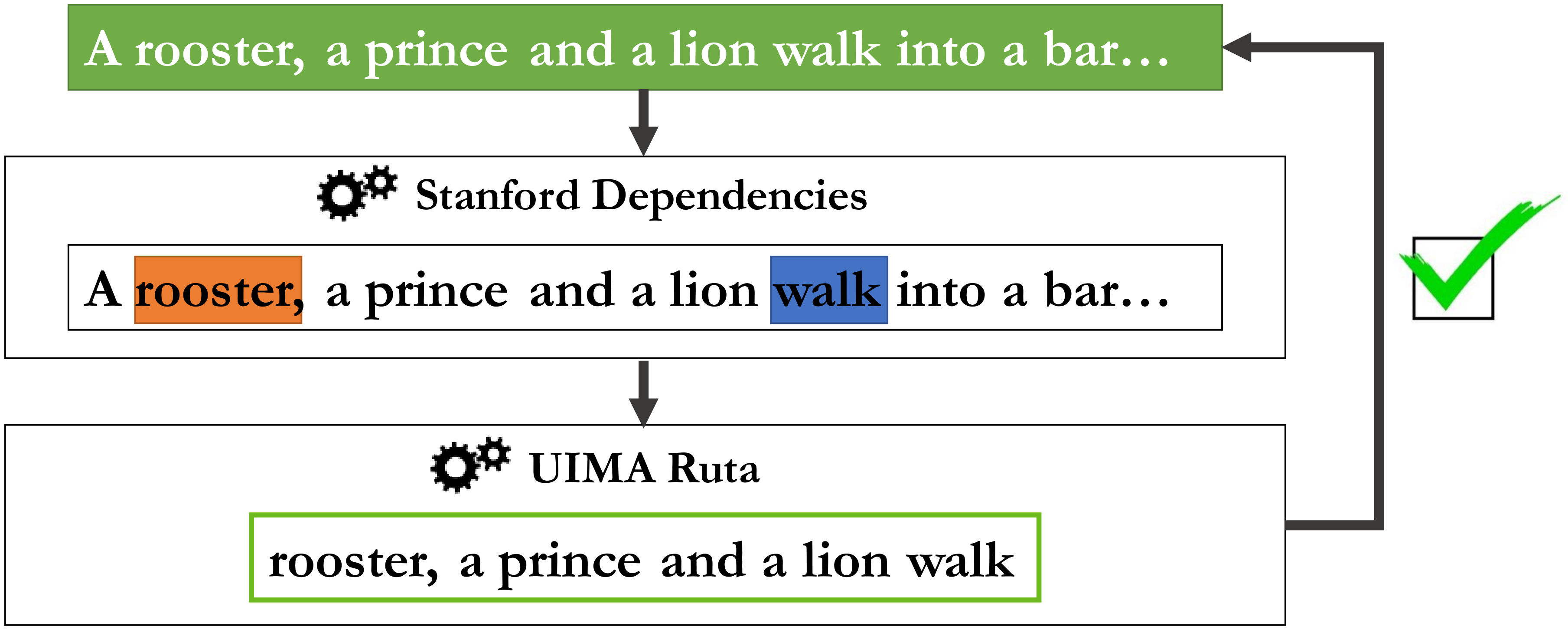
Omission: Hypozeugma

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)



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)



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.

$\boxed{\textbf{If I were president, I would cut taxes.}}$

Custom: If-conditional 2

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

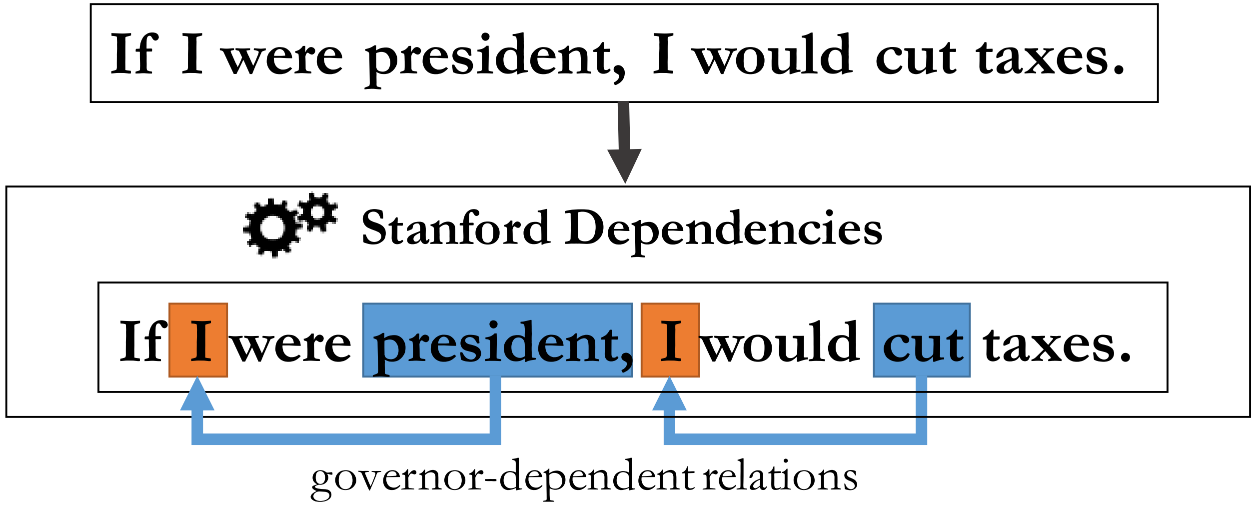
If I were president, I would cut taxes.

Stanford Dependencies

If I were president, I would cut taxes.

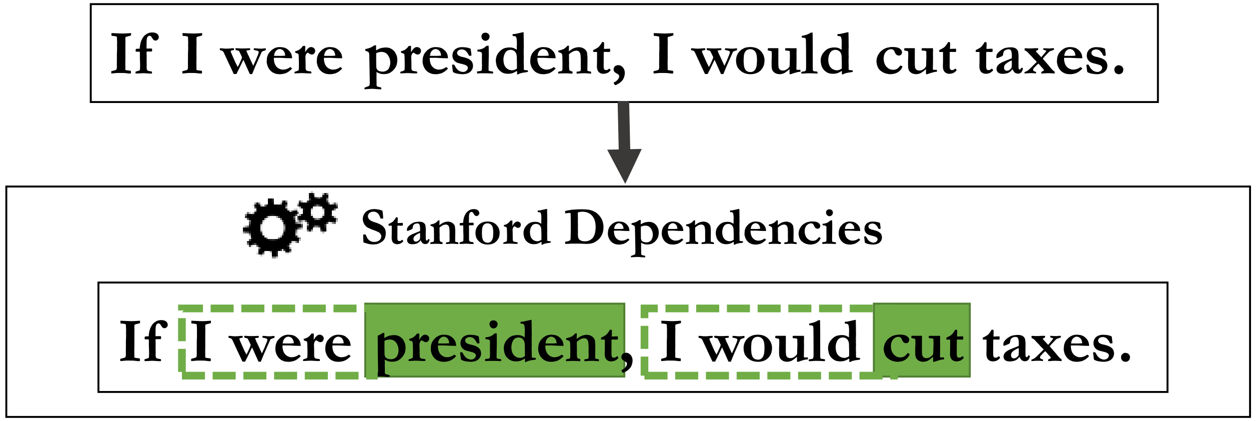
Custom: If-conditional 2

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



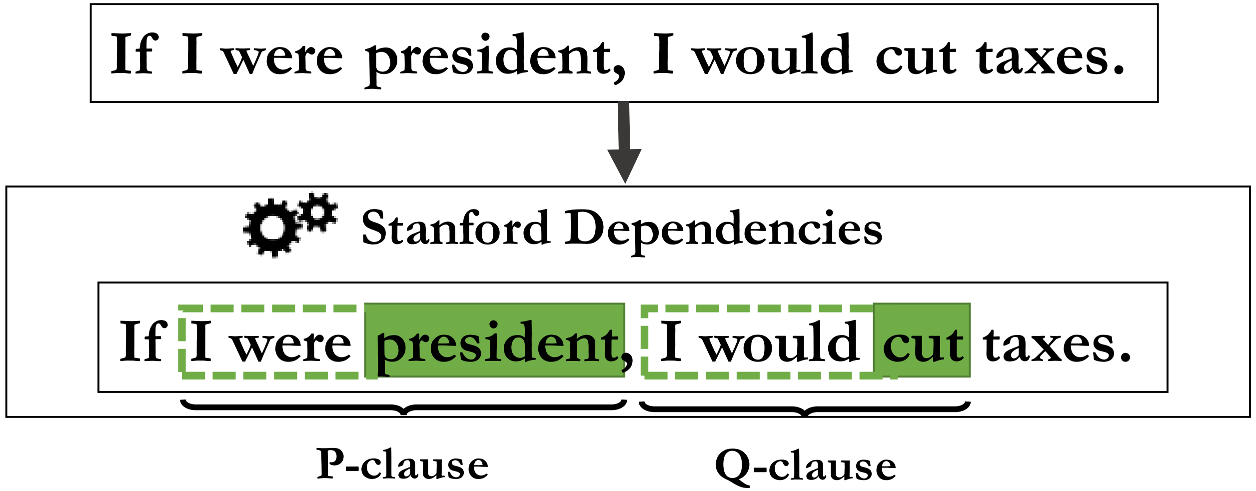
Custom: If-conditional 2

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



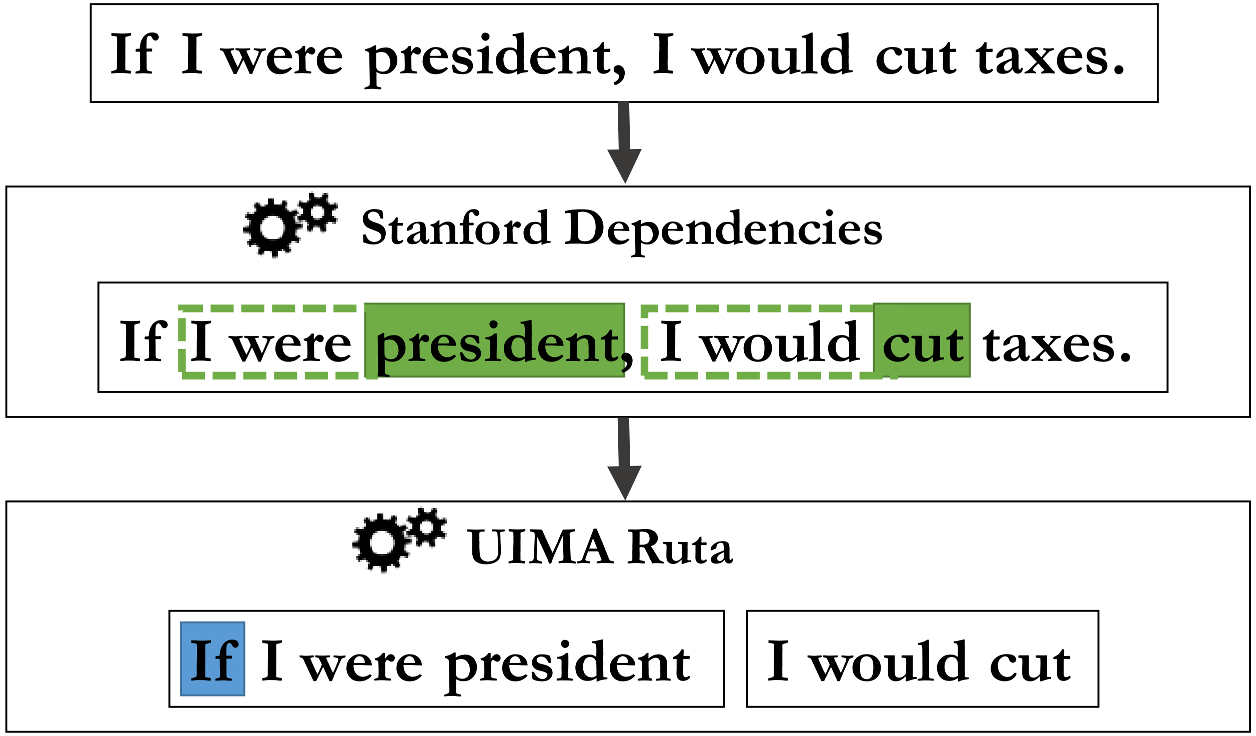
Custom: If-conditional 2

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



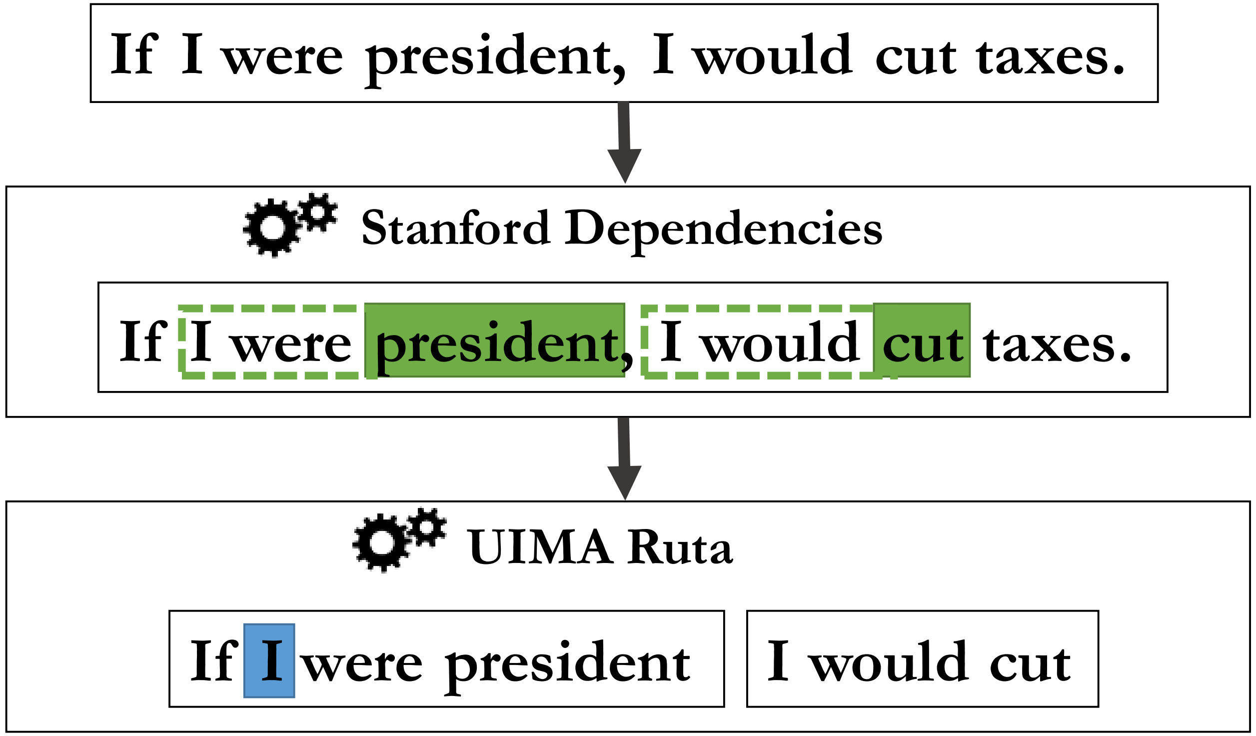
Custom: If-conditional 2

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



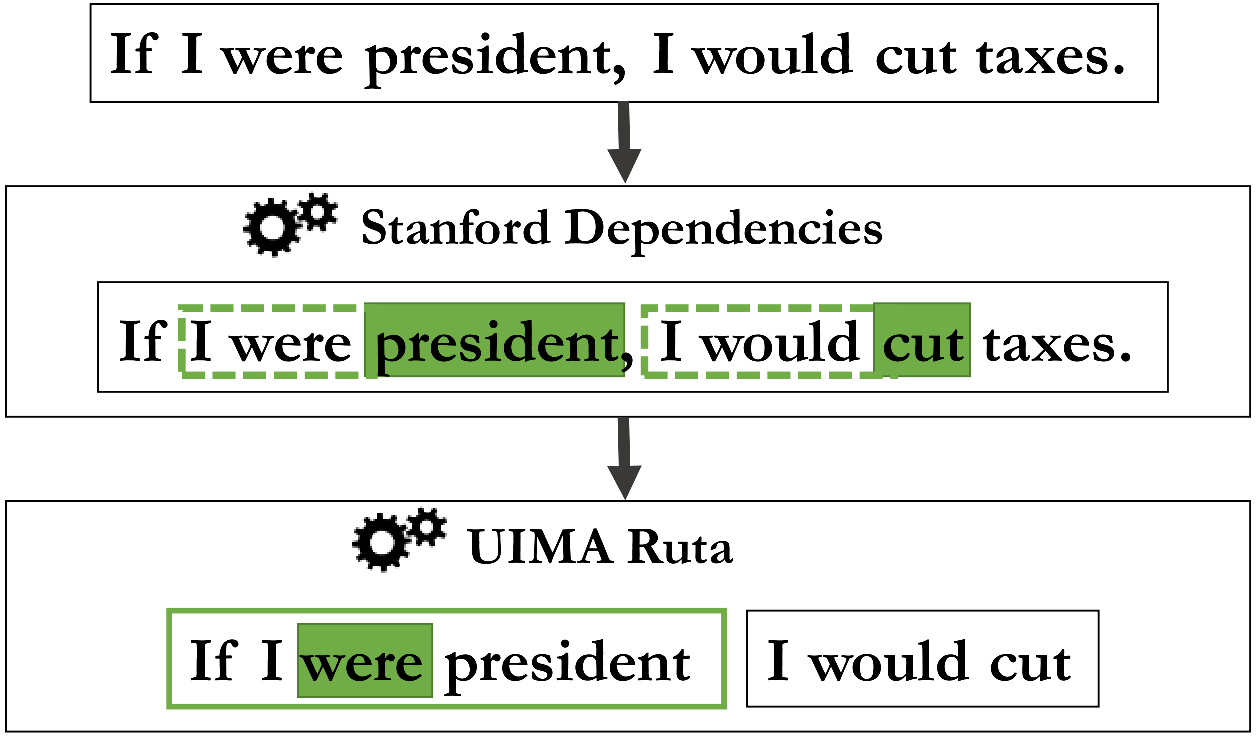
Custom: If-conditional 2

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



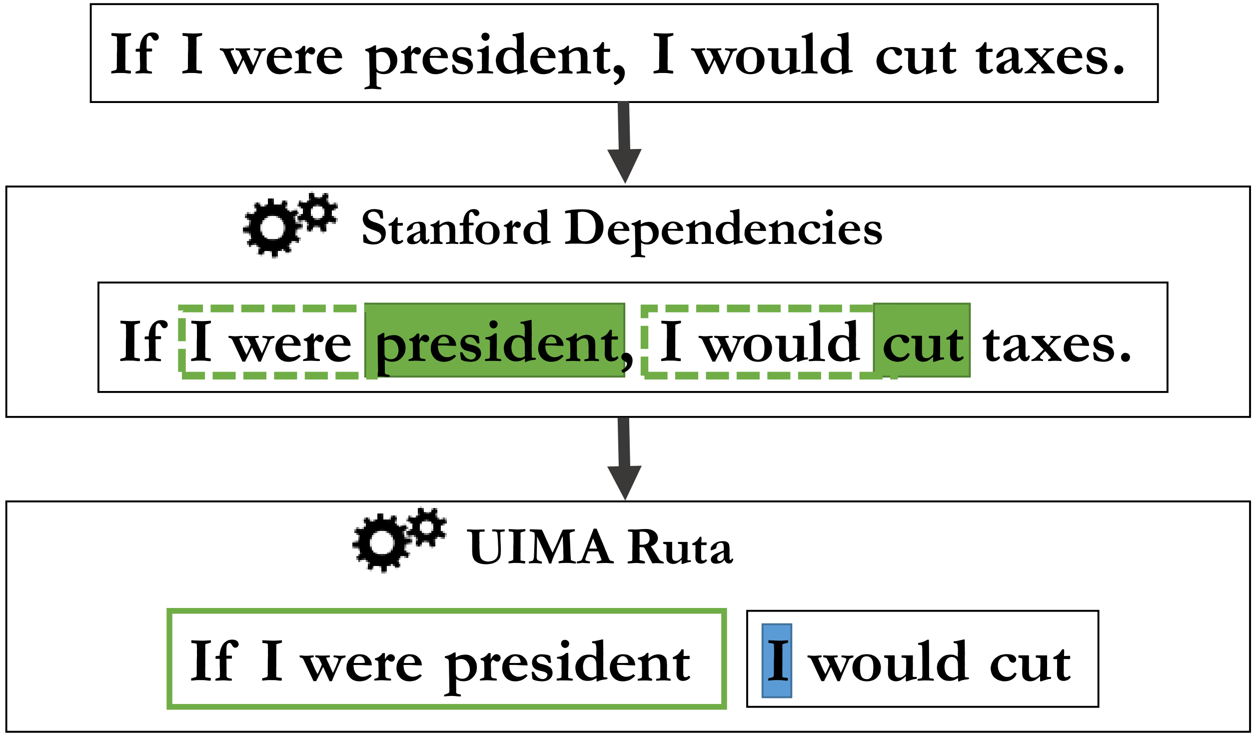
Custom: If-conditional 2

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



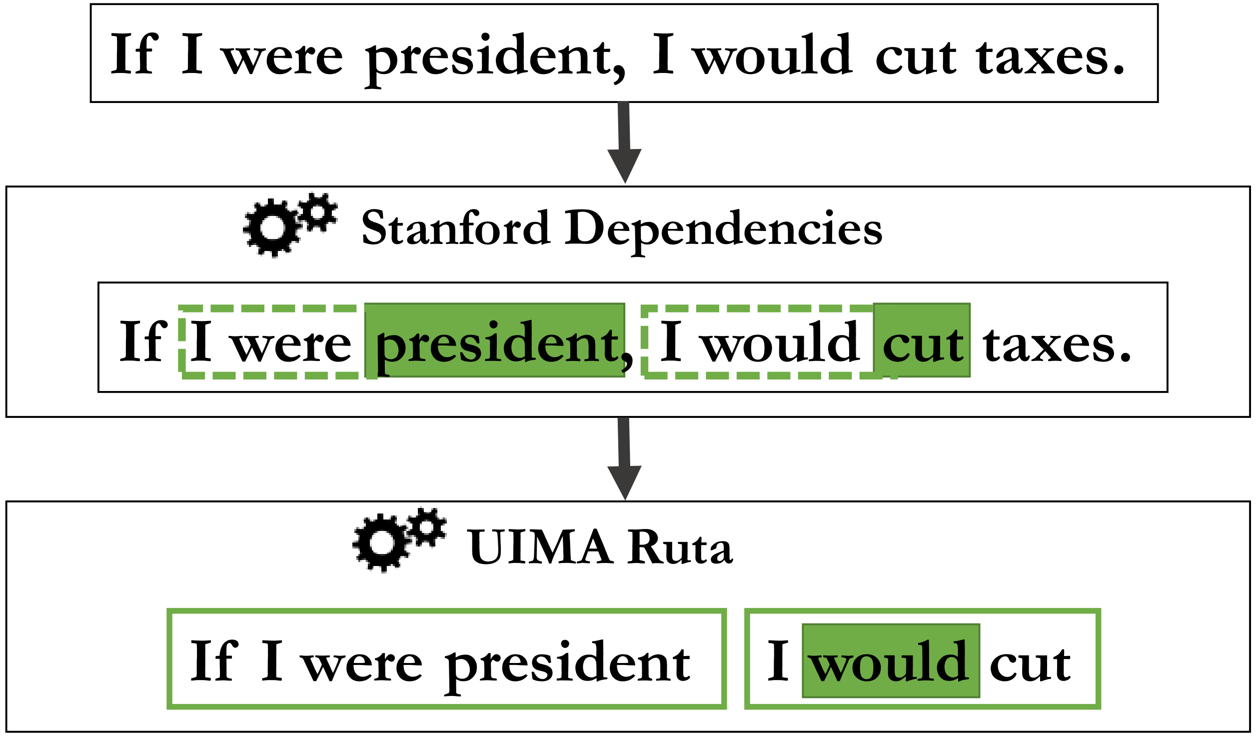
Custom: If-conditional 2

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



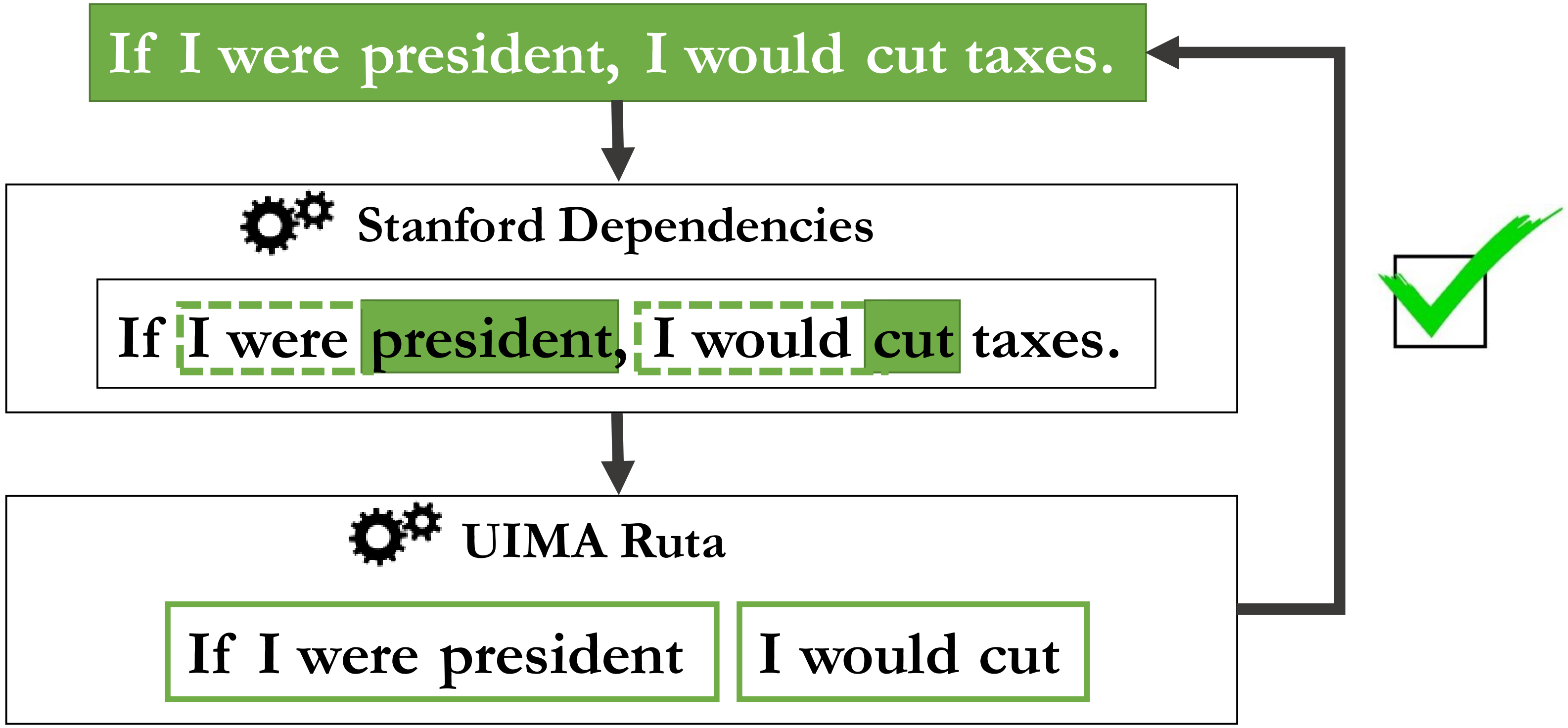
Custom: If-conditional 2

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

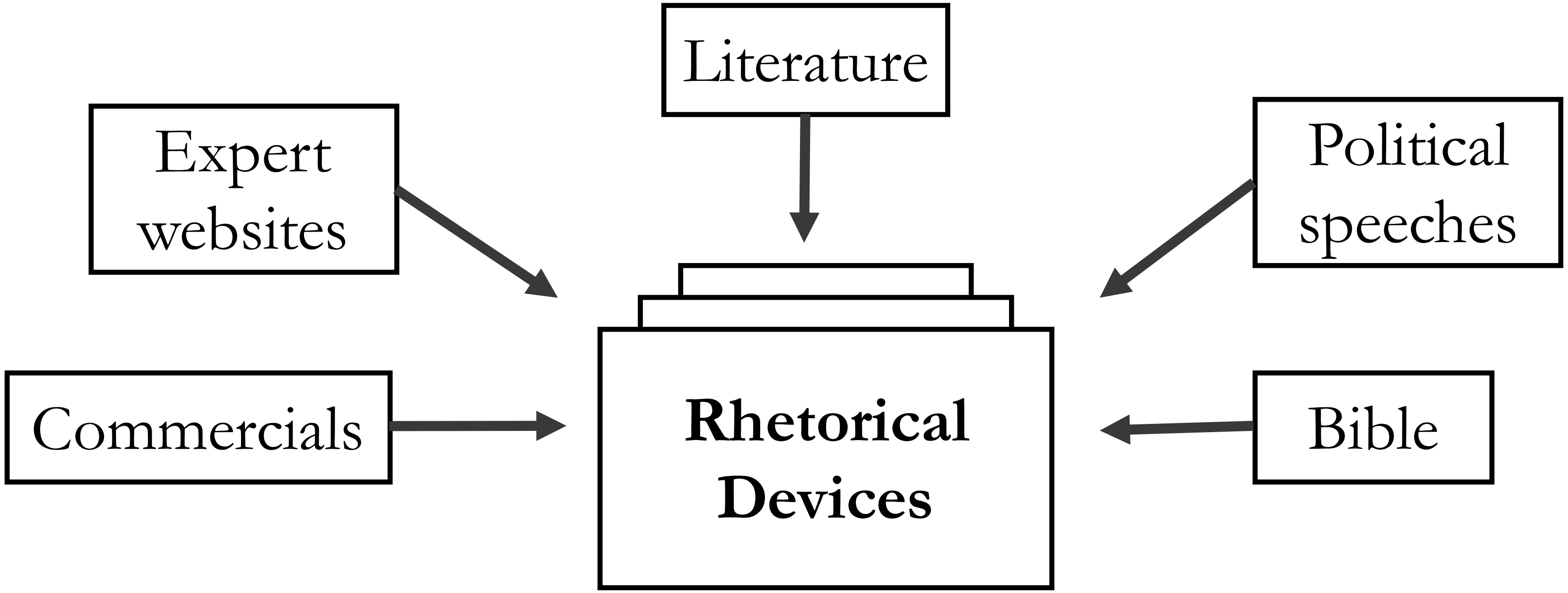


Custom: If-conditional 2

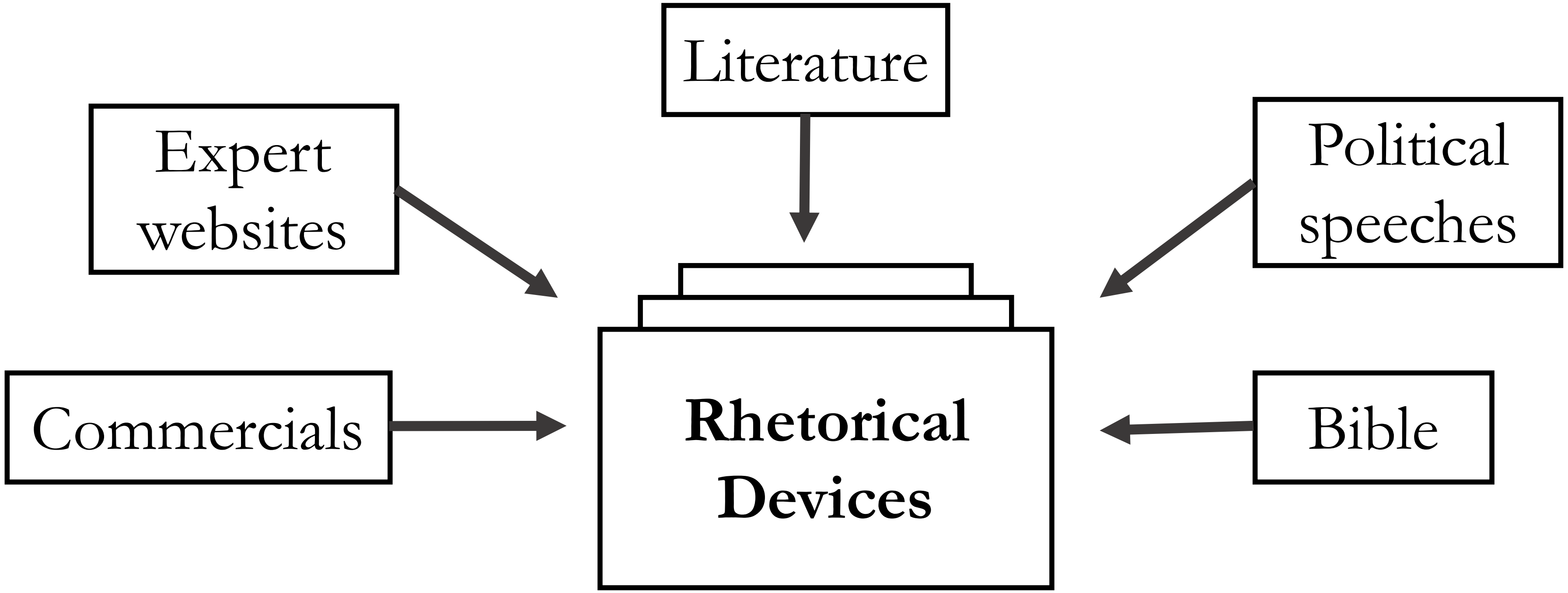
If-conditional 2 - expresses consequences that are totally unrealistic or will not likely happen in the future.



Evaluation dataset



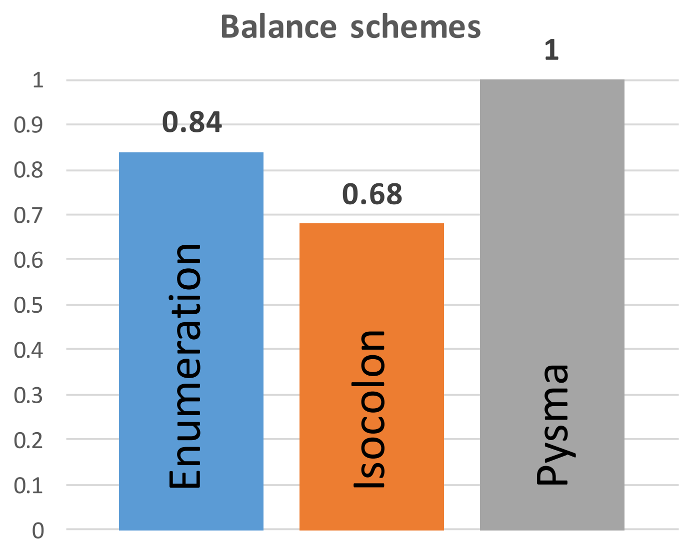
Evaluation dataset

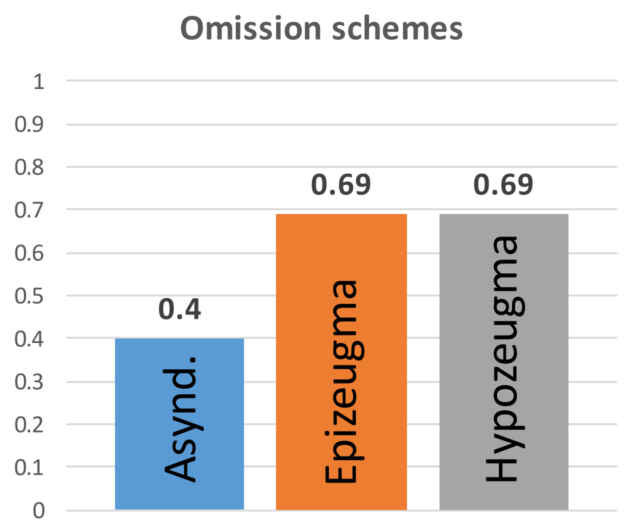


Evaluation measures

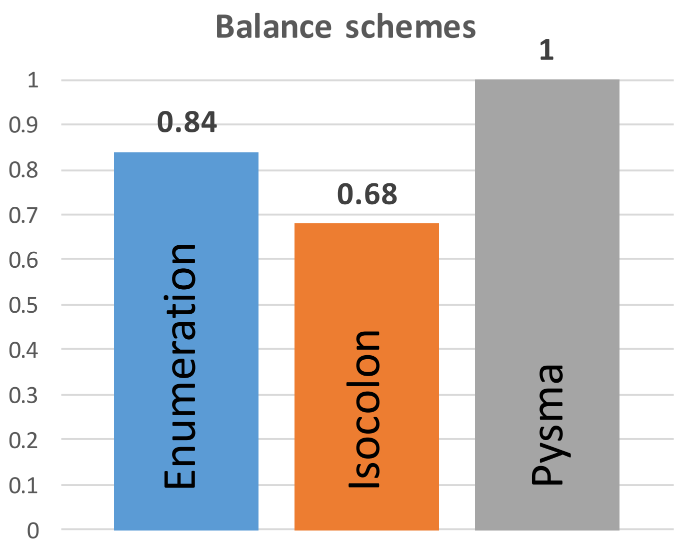
$$Precision = \frac{tp}{tp + fp} \qquad Recall = \frac{tp}{tp + fn} \qquad F1\ score = 2 \cdot \frac{precision \cdot recall}{precision + recall}$$

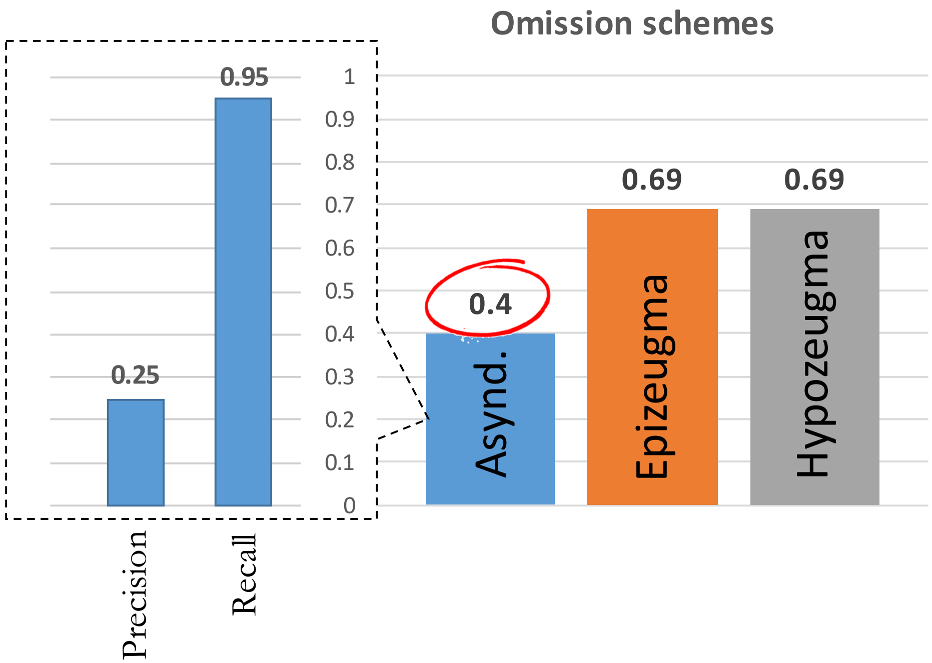
Evaluation Results



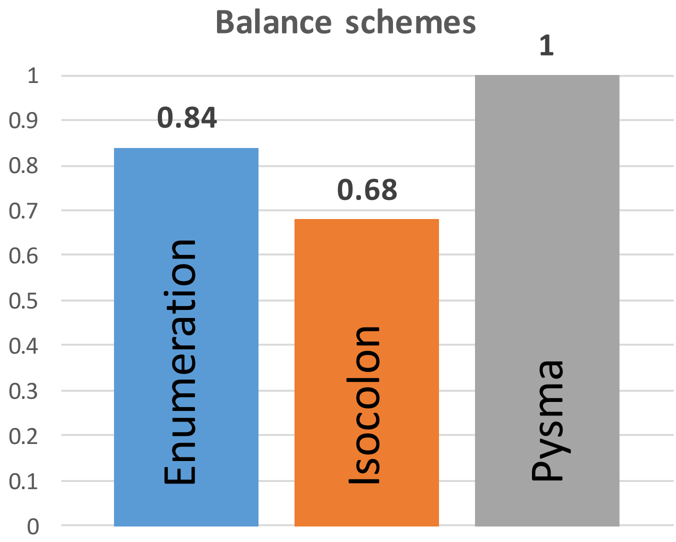


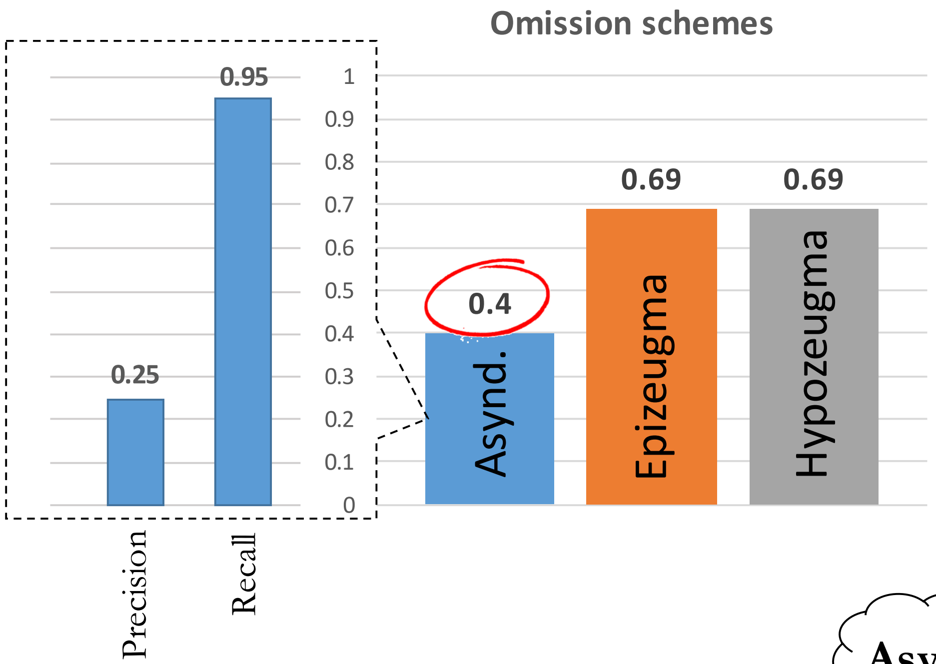
Evaluation Results

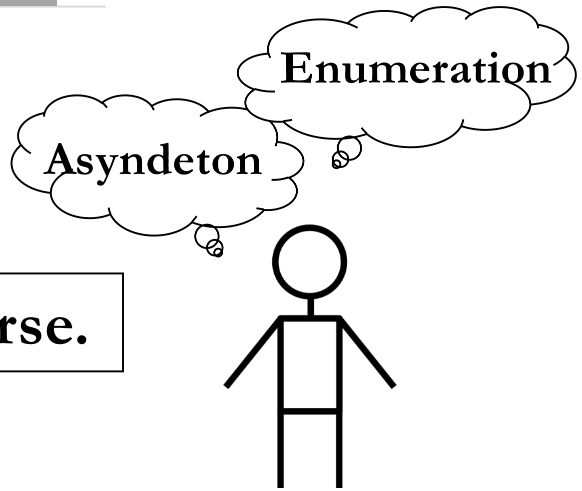




Evaluation Results

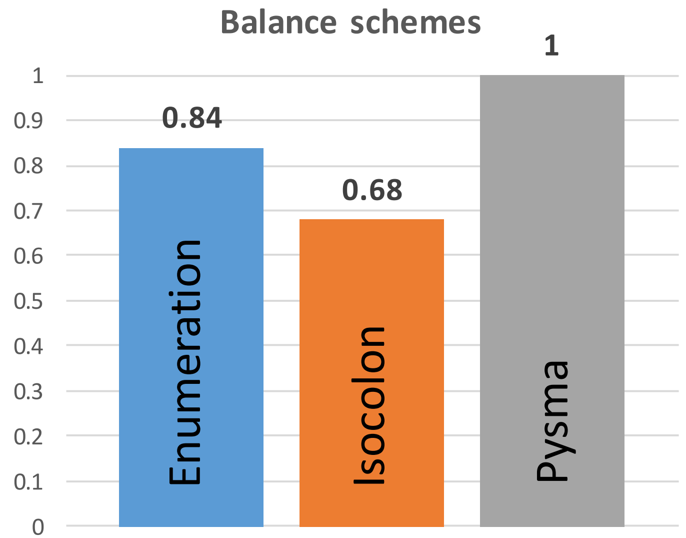


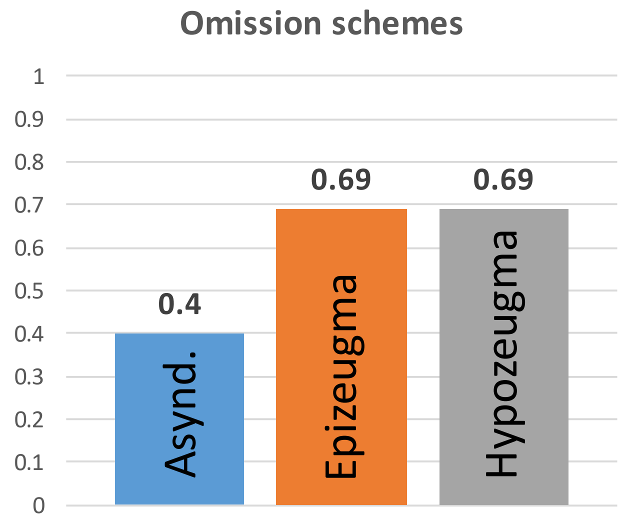


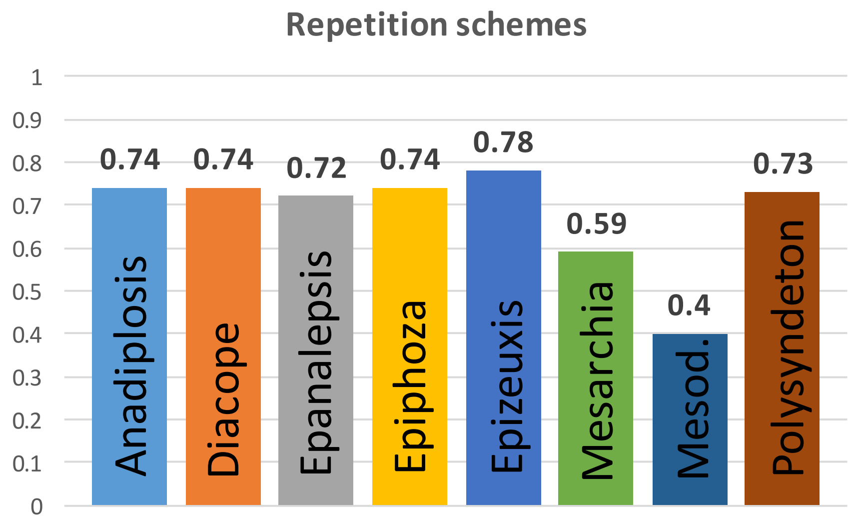


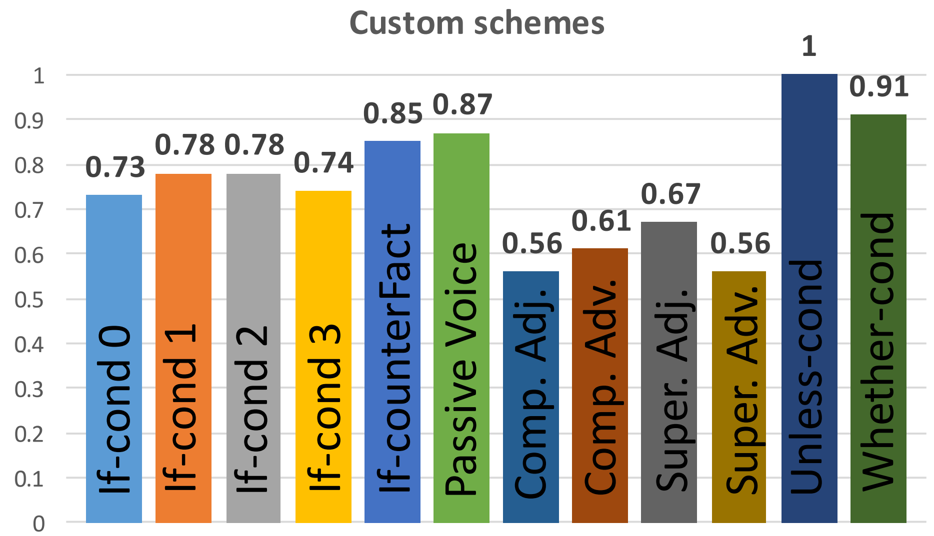
Old farmer had a pig, a dog, a cow and a horse.

Evaluation Results

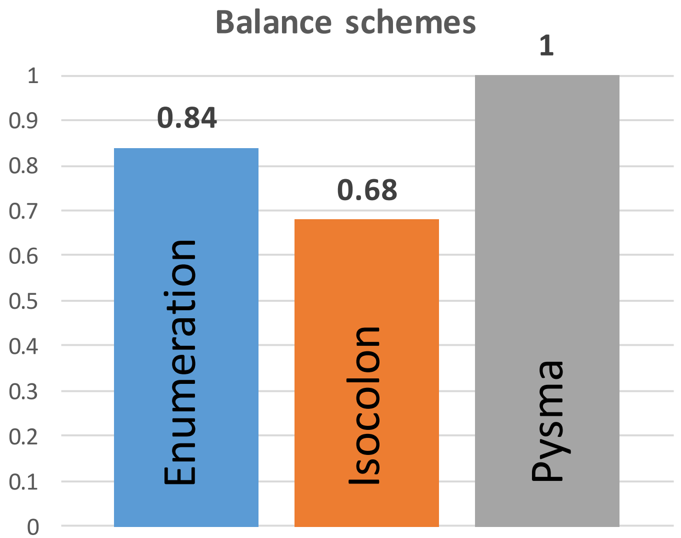


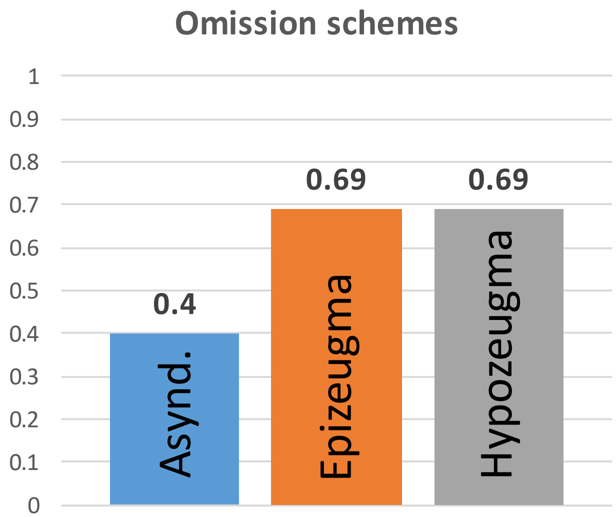


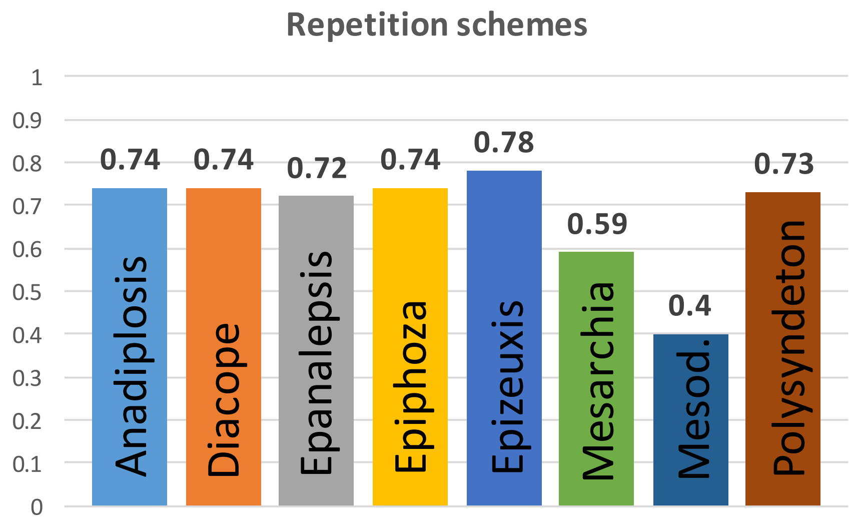


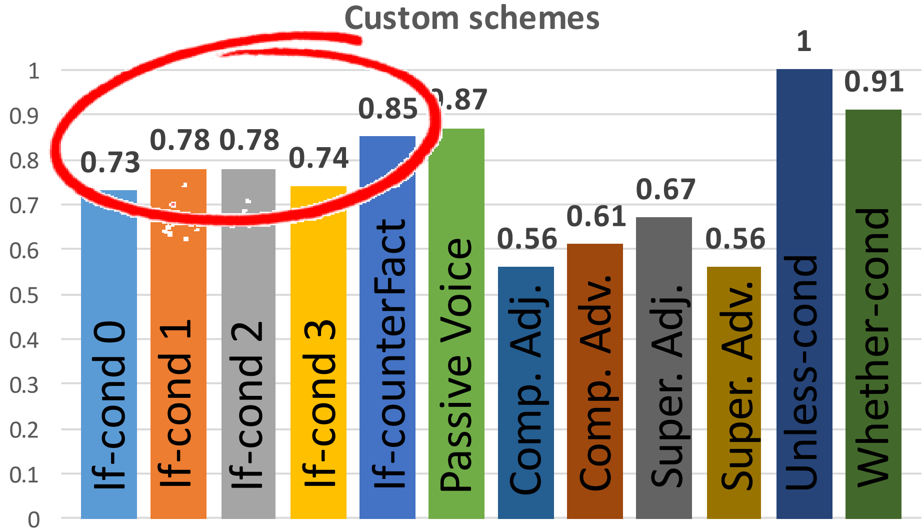


Evaluation Results

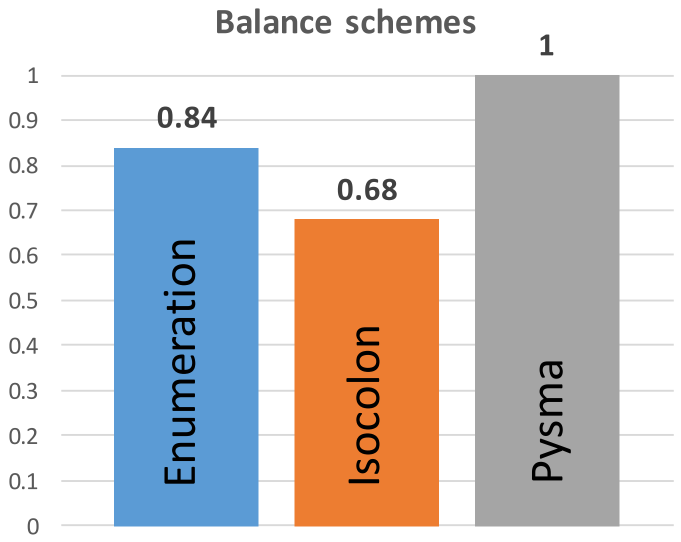


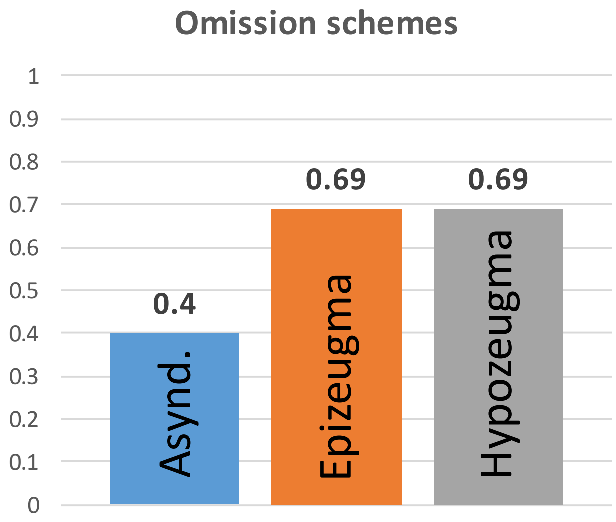


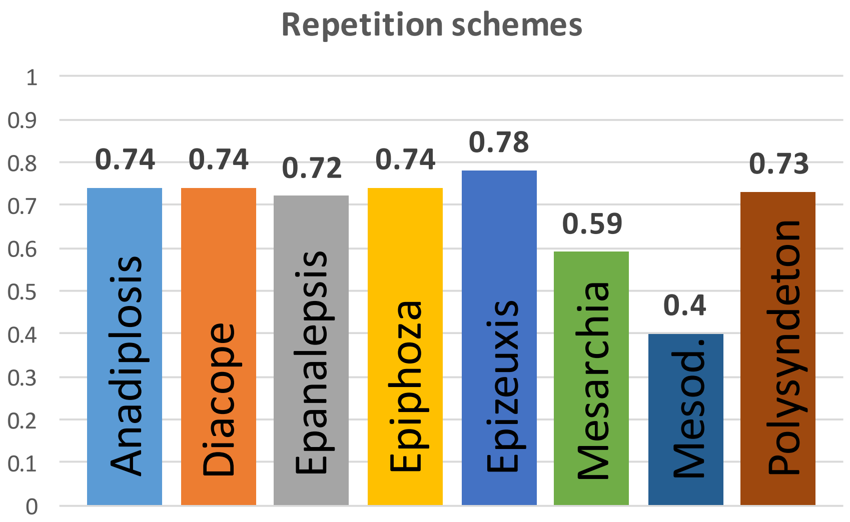


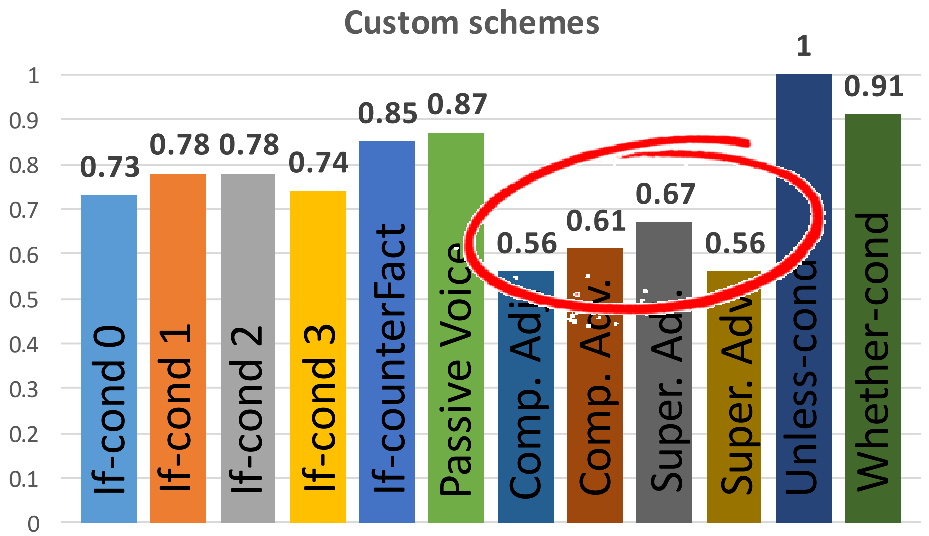


Evaluation Results F1-Score

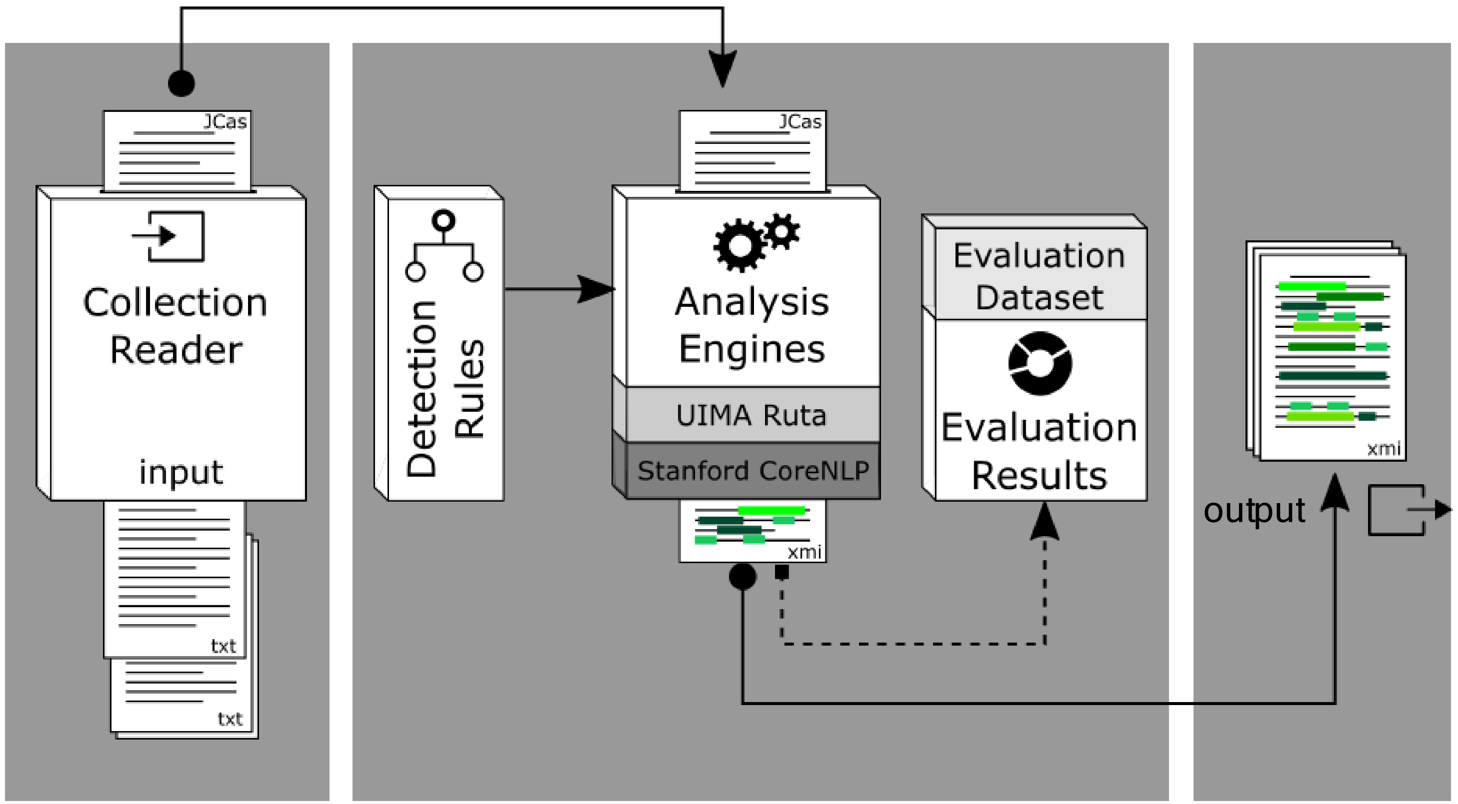






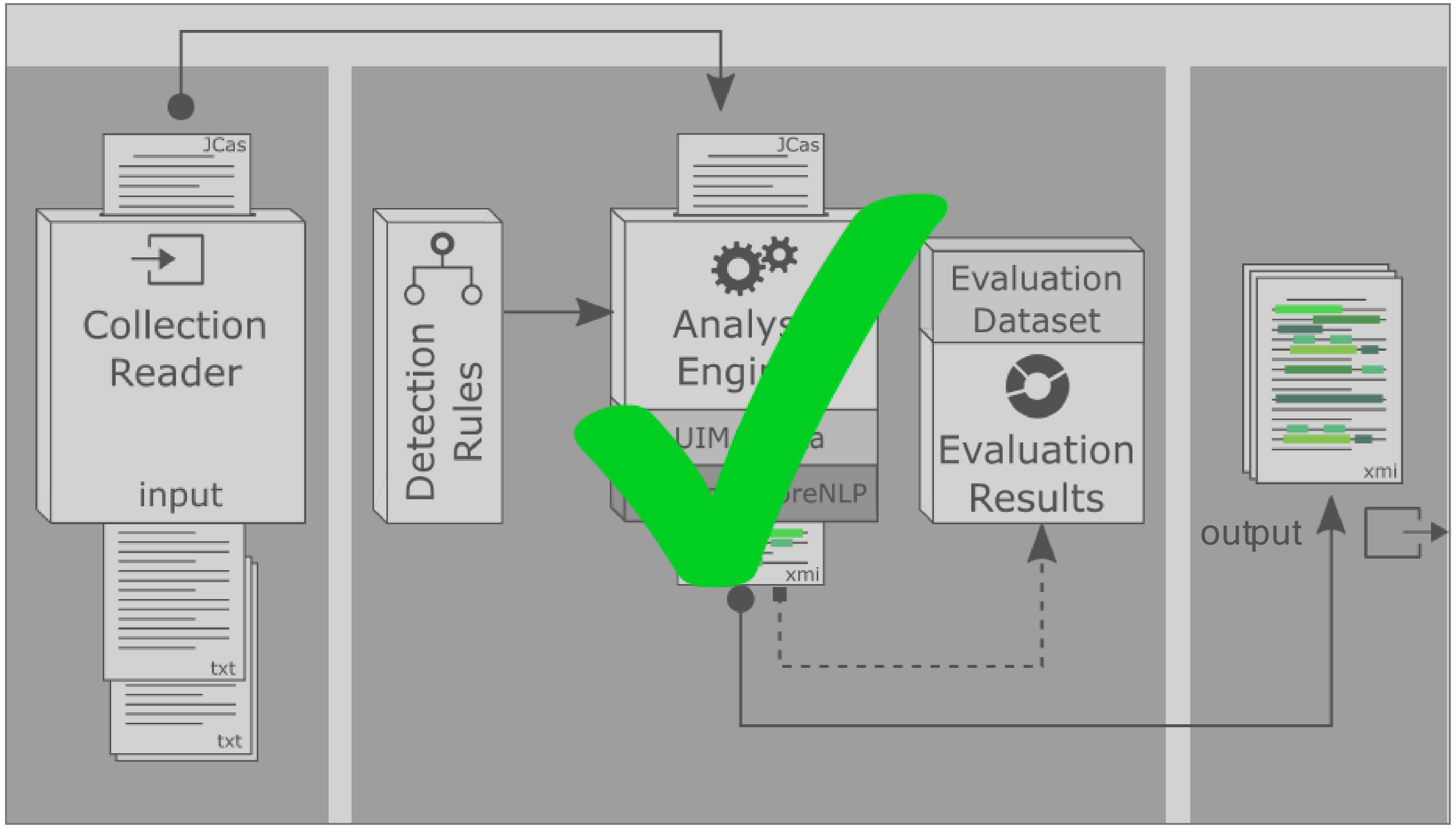


Pipeline



JCas
Collection
Reader
input
txt
txt
Detection
Rules
JCas
Analysis
Engines
UIMA Ruta
Stanford CoreNLP
xmi
Evaluation
Dataset
Evaluation
Results
xmi
output

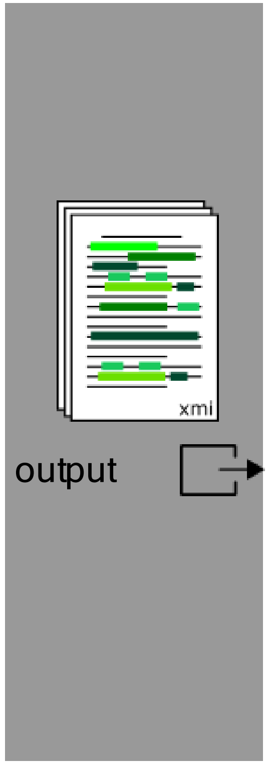
Pipeline



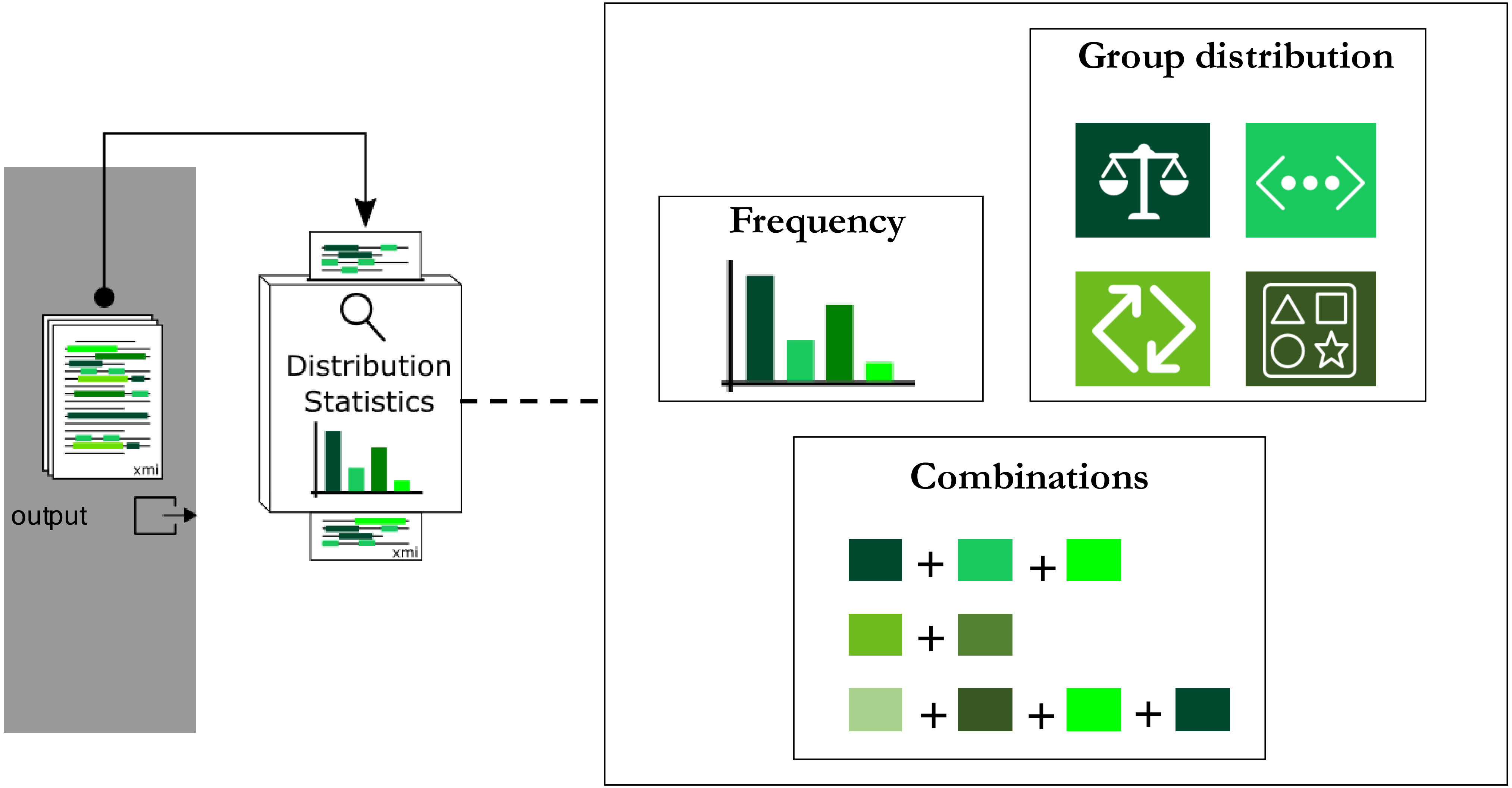
JCas
Collection
Reader
input
txt
txt
Detection
Rules
JCas
Analys
Engi
UIM
a
reNLP
xmi
Evaluation
Dataset
Evaluation
Results
xmi
output

2 Analysis of Rhetorical Devices

Pipeline

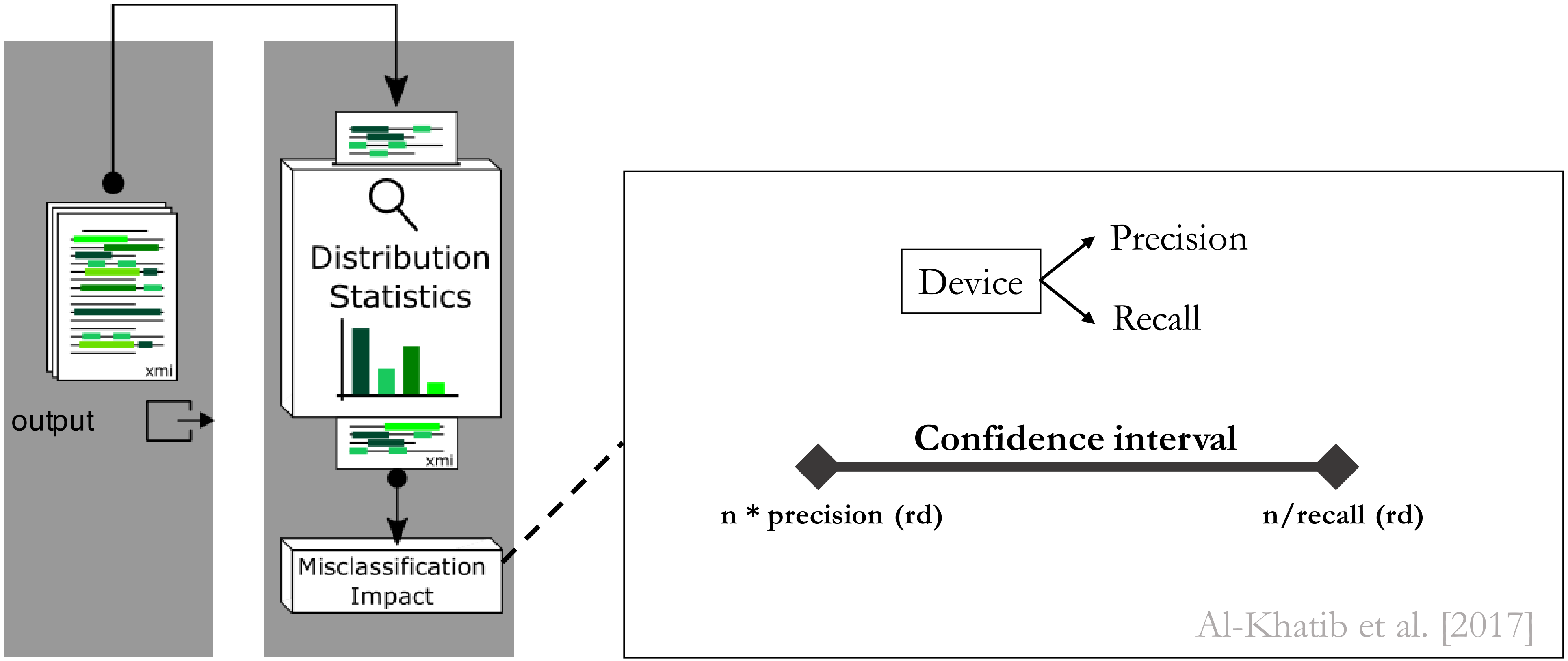


Pipeline



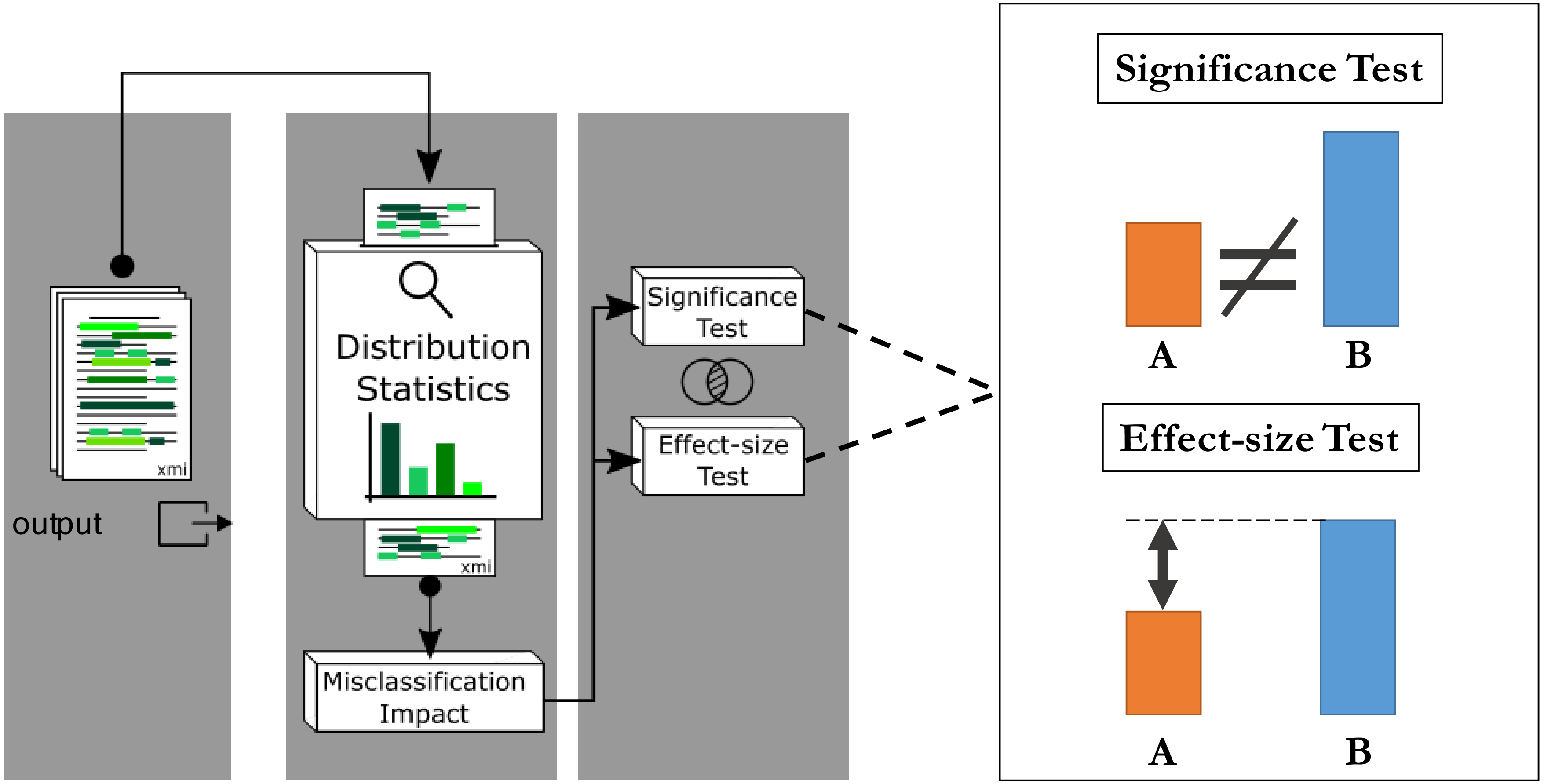
xmi
output
Distribution
Statistics
xmi
Frequency
Group distribution
Combinations

Pipeline



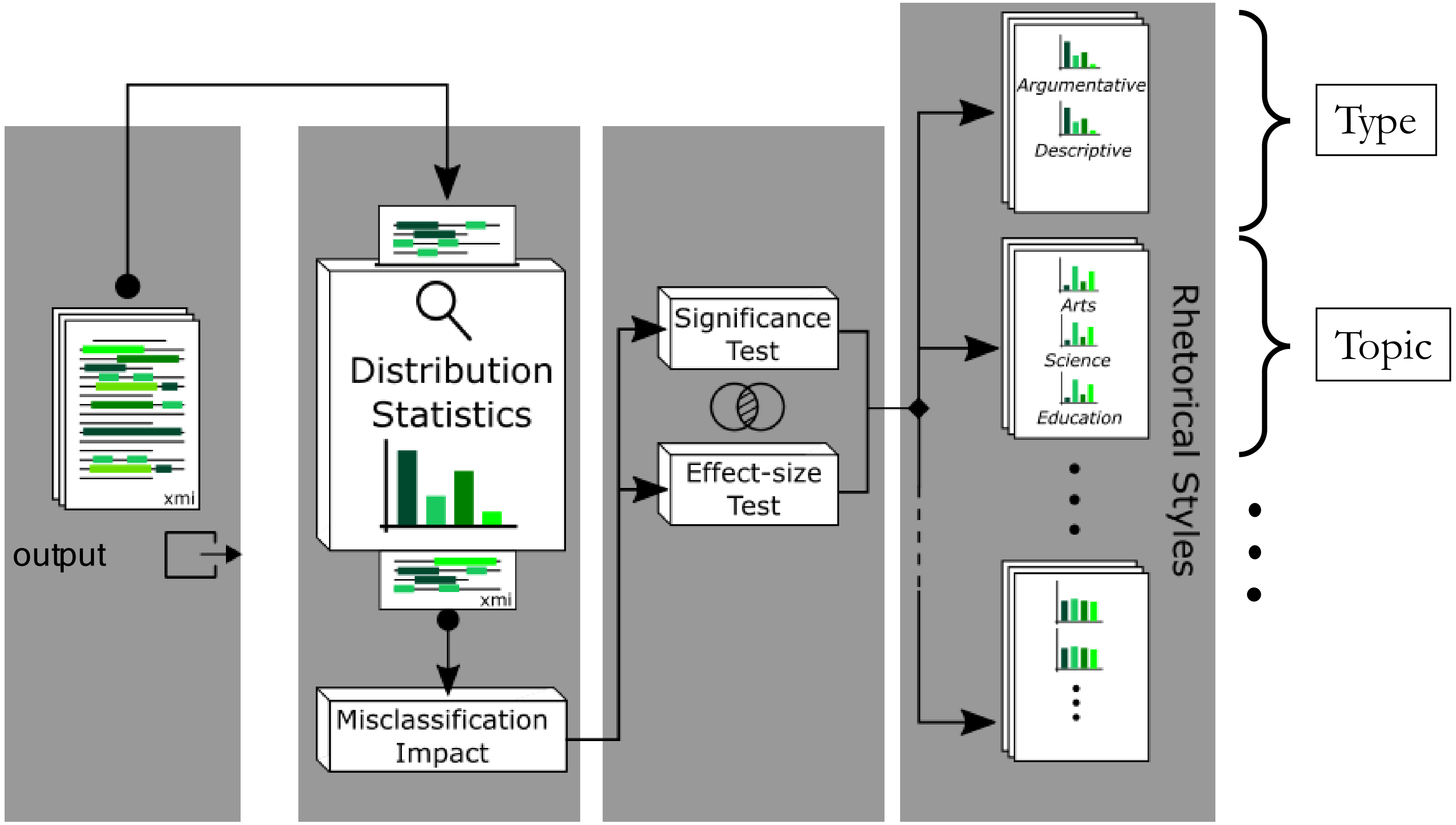
xmi
output
Distribution Statistics
xmi
Misclassification Impact
Device
Precision
Recall
Confidence interval
n * precision (rd)
n/recall (rd)
Al-Khatib et al. [2017]

Pipeline



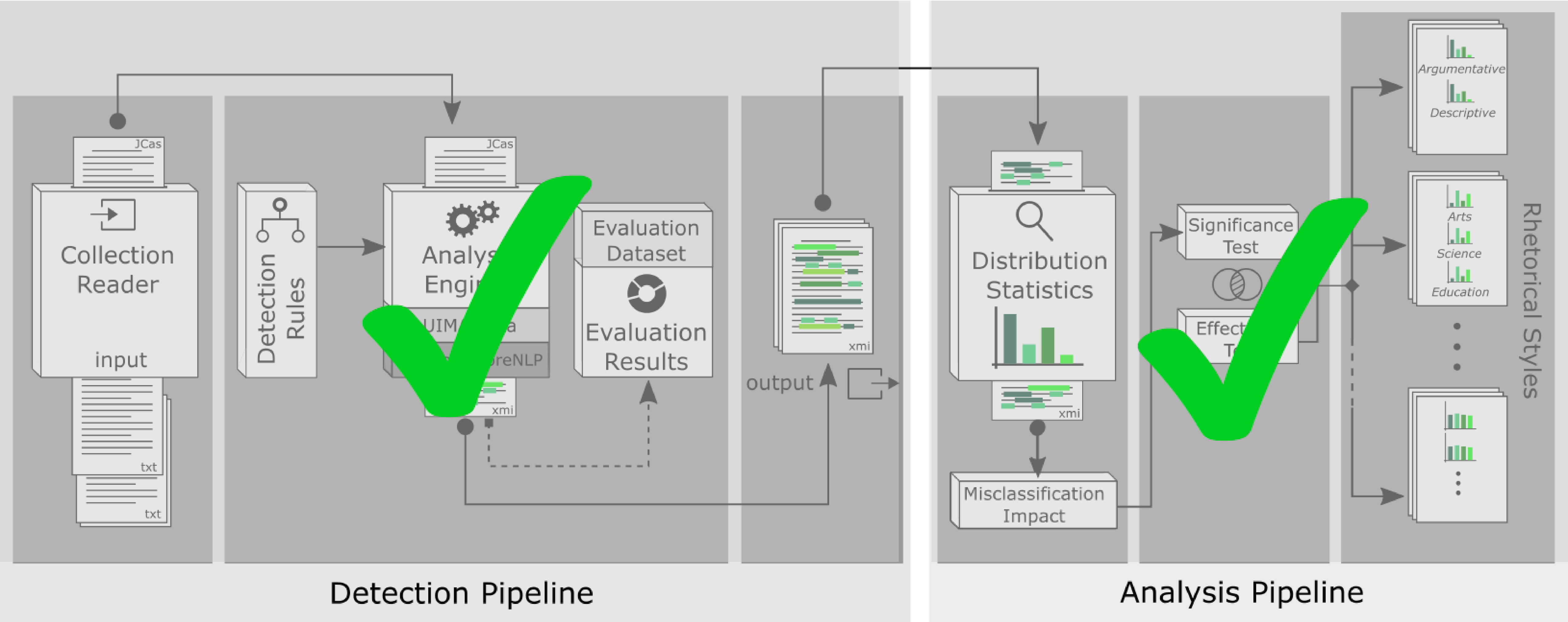
xmi
output
Distribution Statistics
xmi
Misclassification Impact
Significance Test
Effect-size Test
Significance Test
A
B
Effect-size Test
A
B

Pipeline



xmi
output
Distribution Statistics
xmi
Misclassification Impact
Significance Test
Effect-size Test
Argumentative
Descriptive
Arts
Science
Education
Rhetorical Styles
Type
Topic

Pipeline



JCas
Collection Reader
input
txt
txt
Detection Rules
JCas
Analys
Engi
UIM
reNLP
xmi
Evaluation Dataset
Evaluation Results
xmi
output
xmi
Distribution Statistics
xmi
Misclassification Impact
Significance Test
Effect
T
Argumentative
Descriptive
Arts
Science
Education
Rhetorical Styles
Detection Pipeline
Analysis Pipeline

3 Analysis Experiments

Data Preparation

Experiments: datasets

The New York Times

US Presidential Debates 2016

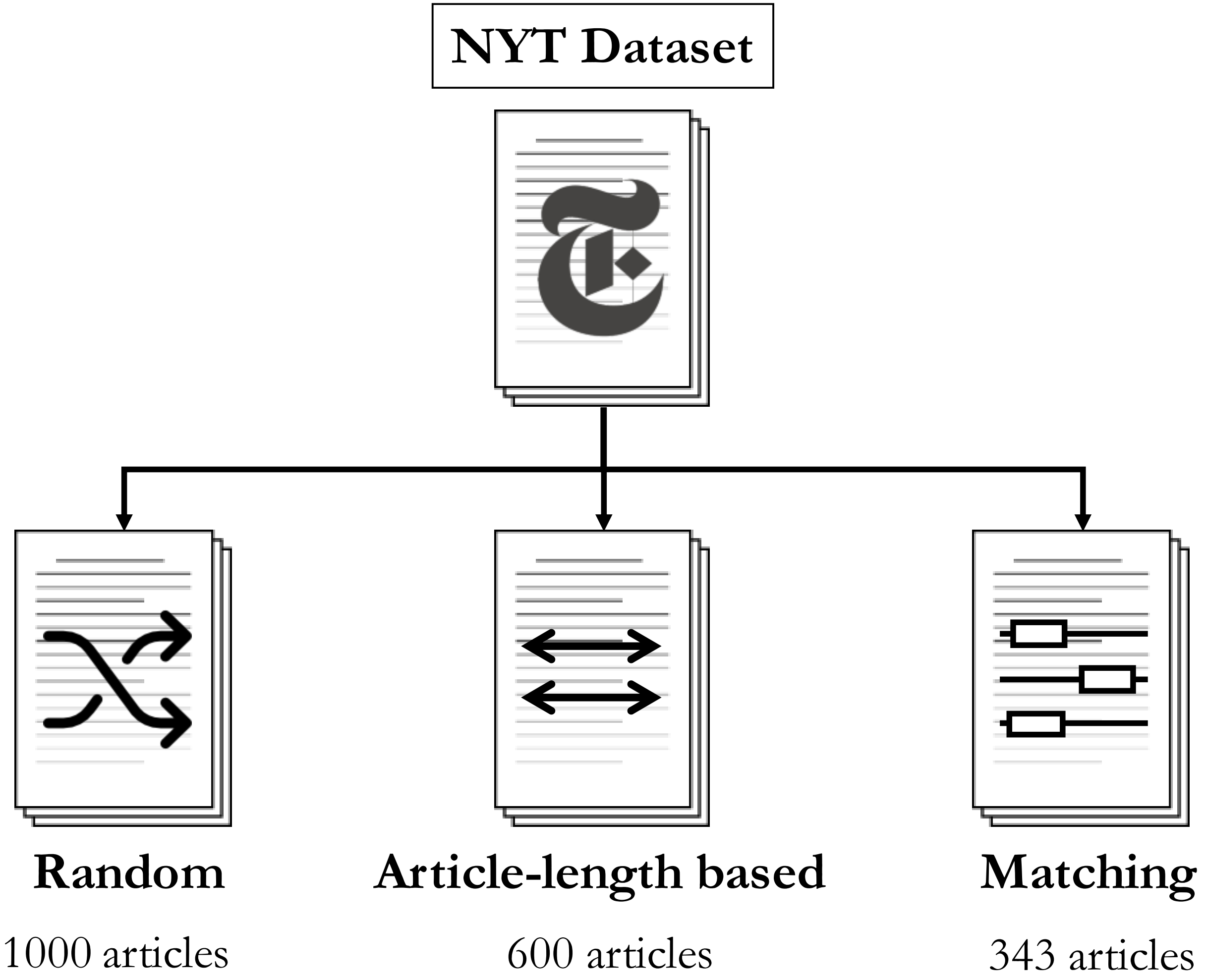
Ben Wiseman [2016]

Data dimensionality

Language	Mode	Communication	Author	Audience
English	Written	Monological	Identity	U.S.

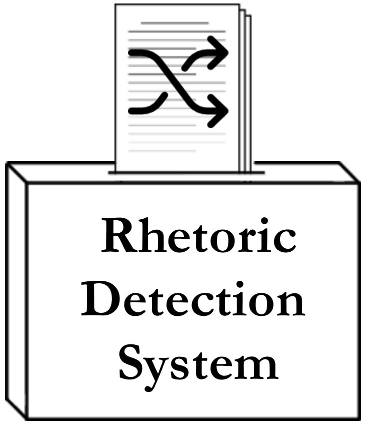
Type	Genre	Topic	Medium
Descriptive	Editorial	Education	Newspaper
Argumentative	Review	Science	Presidential Debates
	Biography	Art	
	Debate	Politics	

NYT Experiment: data subsampling

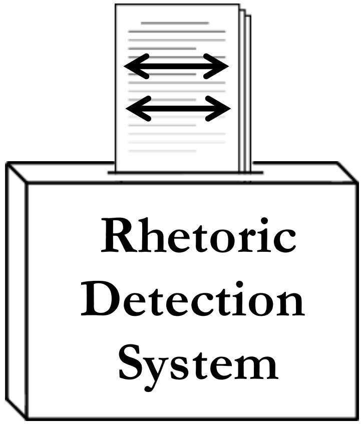


NYT Experiment: Findings

“Random” dataset



“Article-length based” dataset



Articles cover multiple dimensions

Hard to deduce particular styles

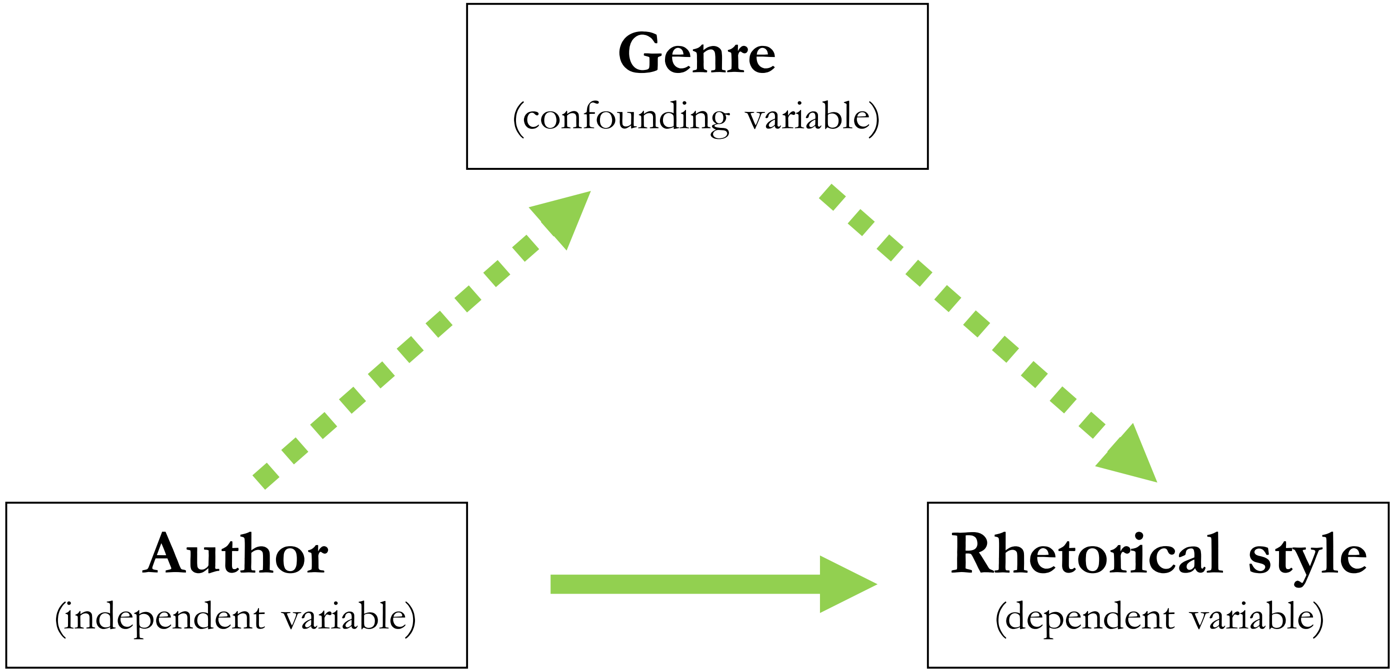
NYT Experiment: Findings

"Random" dataset **"Article-length based" dataset**

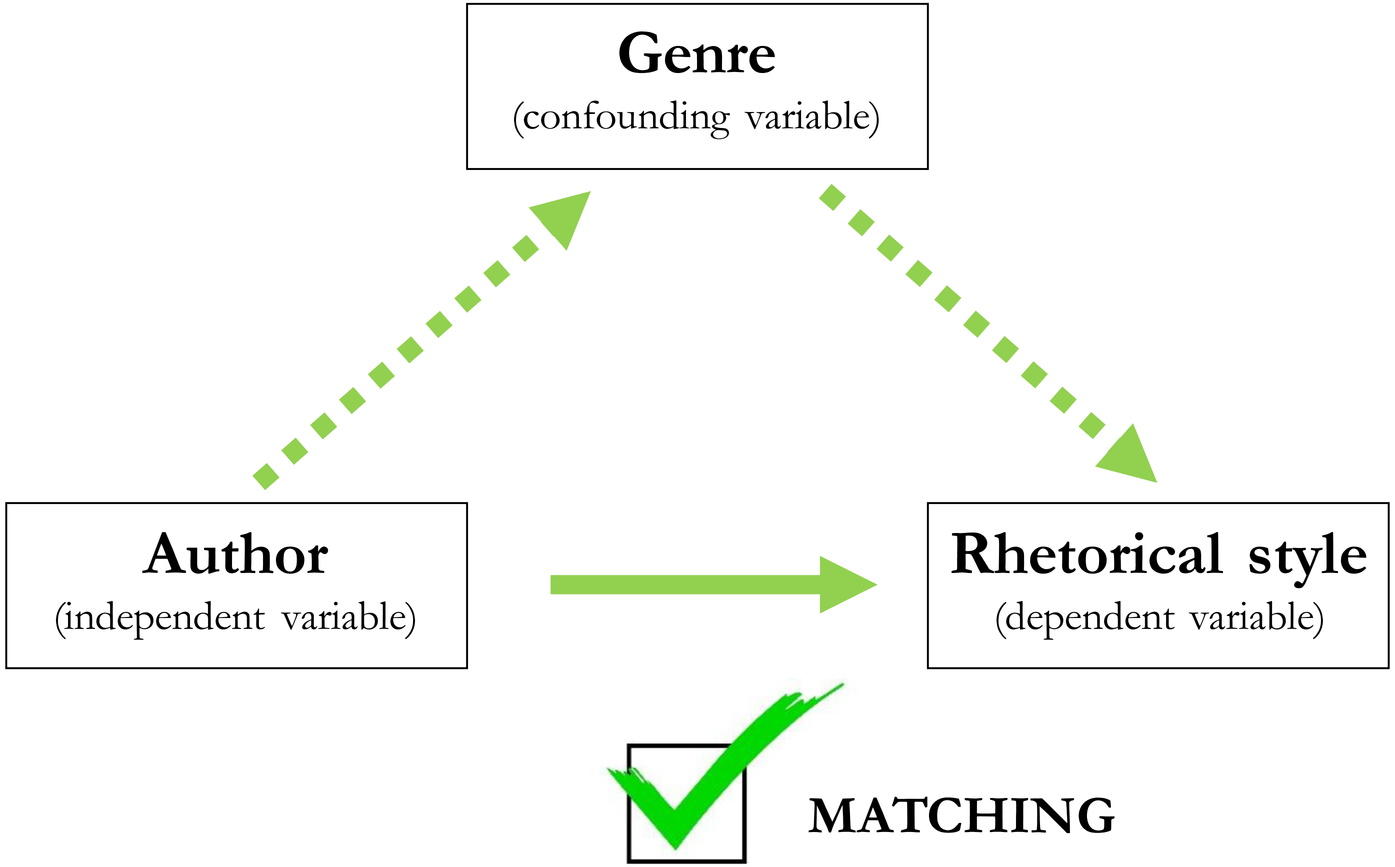
Articles cover multiple dimensions

Hard to deduce particular styles

NYT Experiment: Confounding



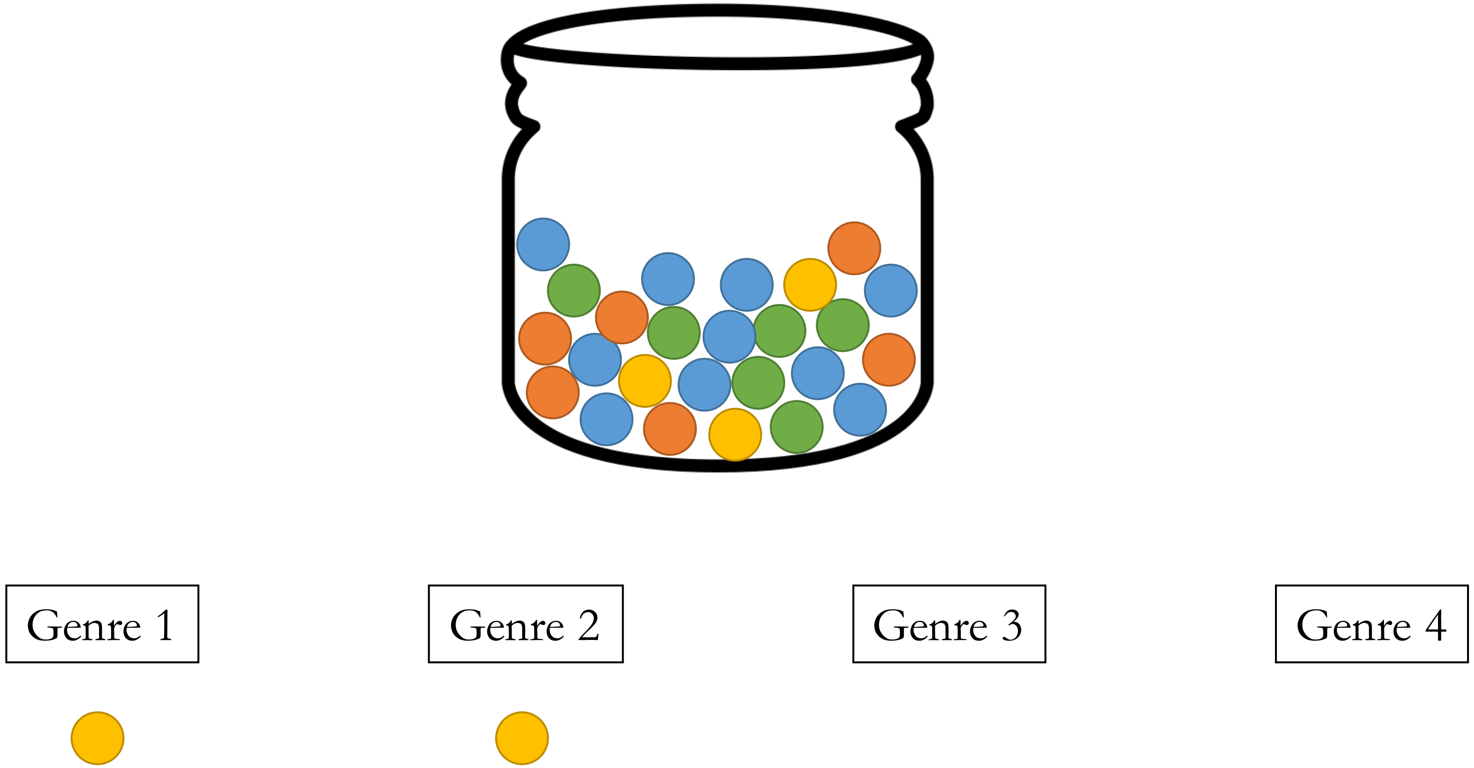
NYT Experiment: Confounding



NYT Experiment: Matching

Genre 1 | Genre 2 | Genre 3 | Genre 4

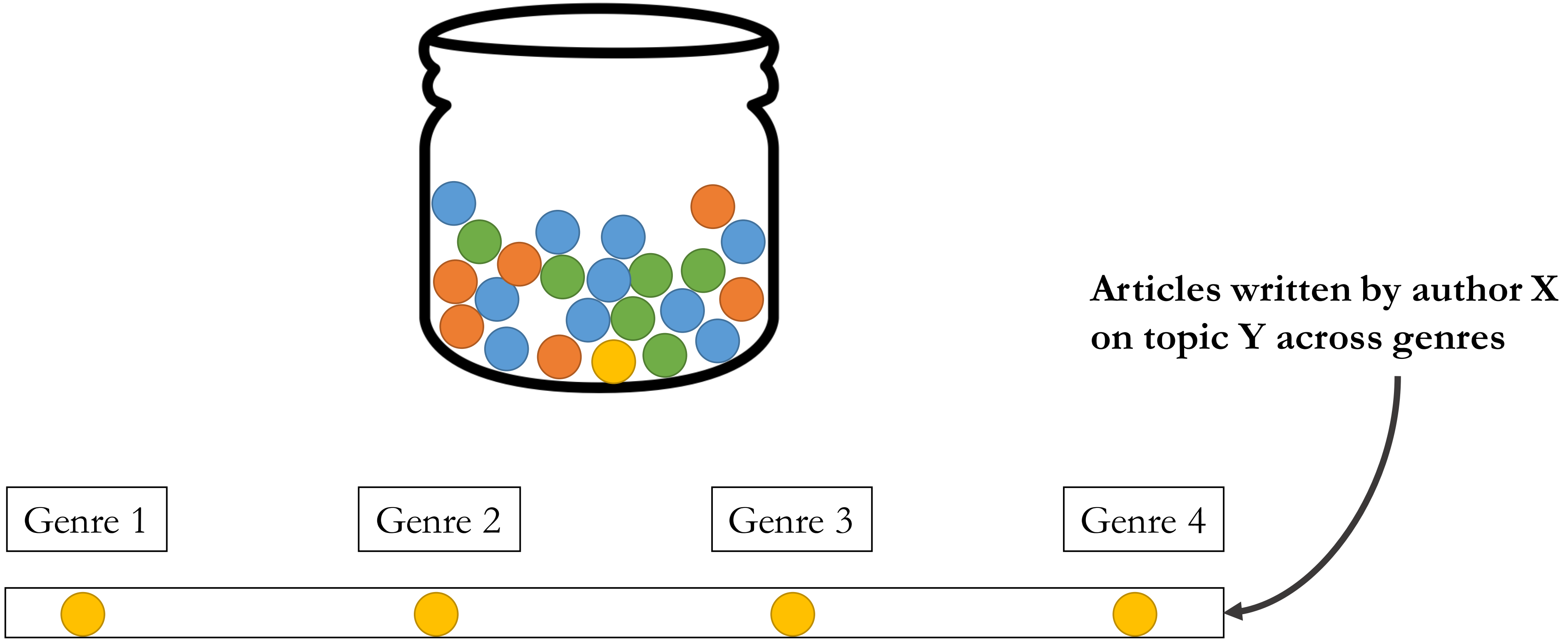
NYT Experiment: Matching



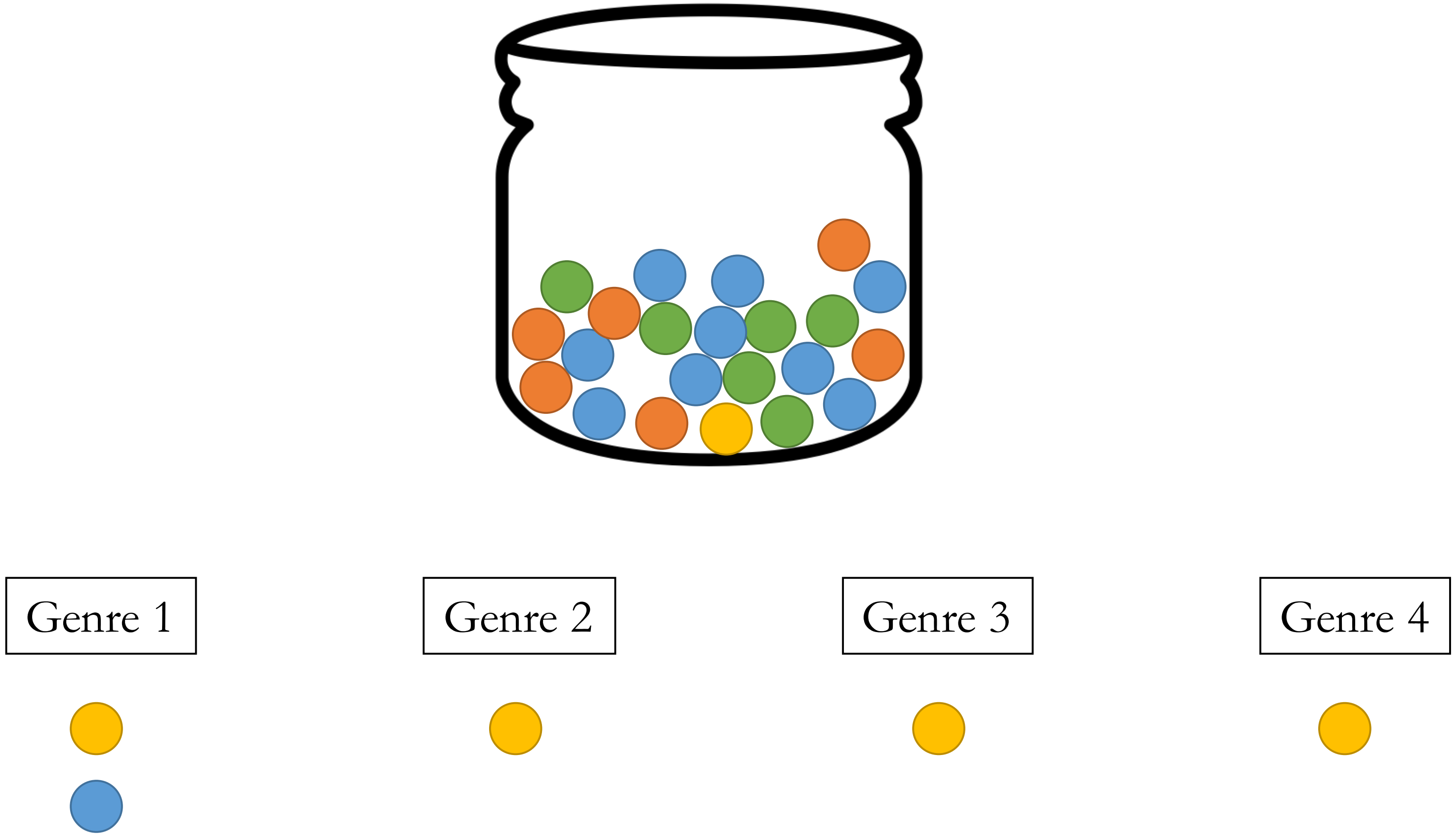
NYT Experiment: Matching

Genre 1 Genre 2 Genre 3 Genre 4

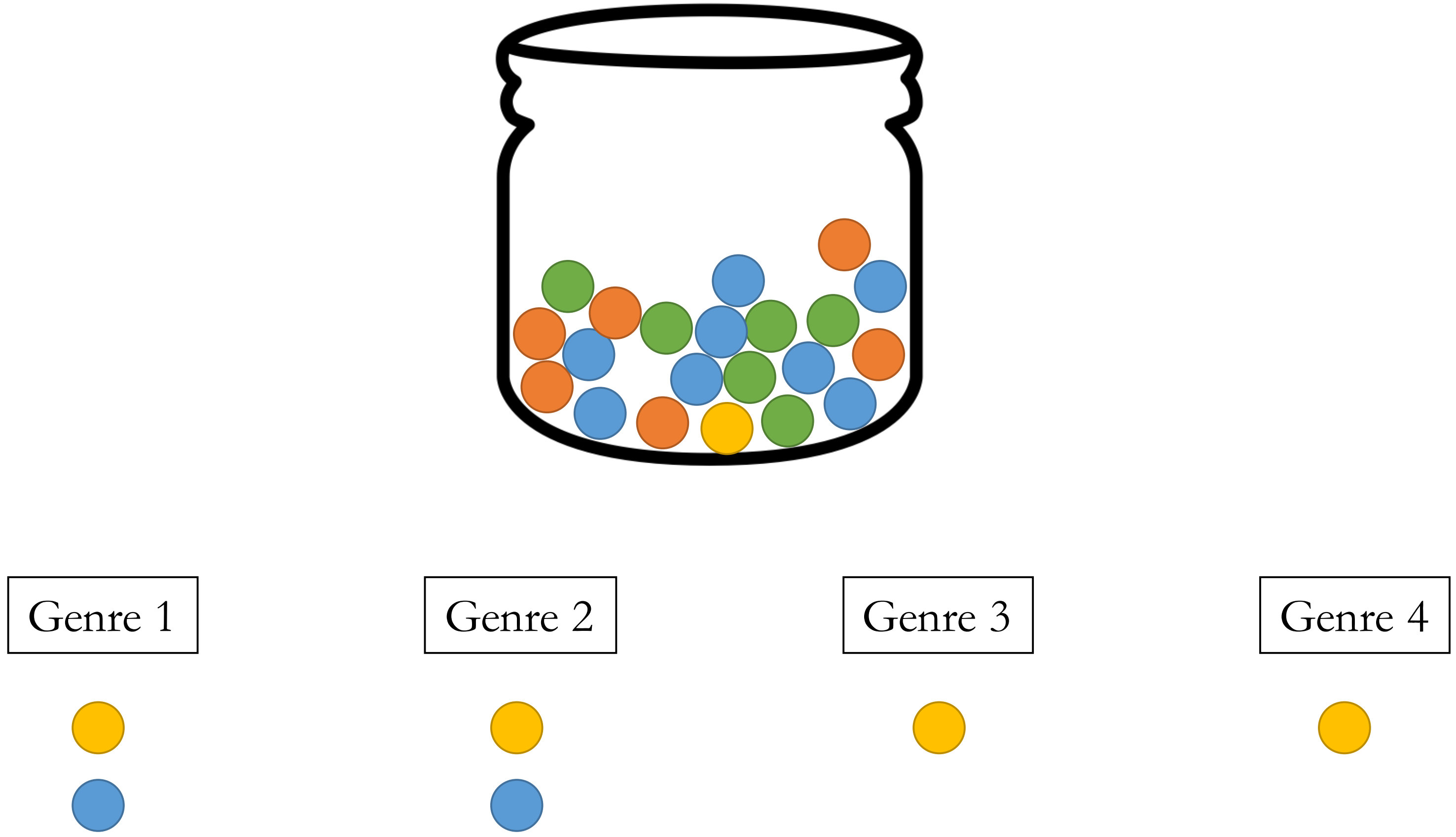
NYT Experiment: Matching



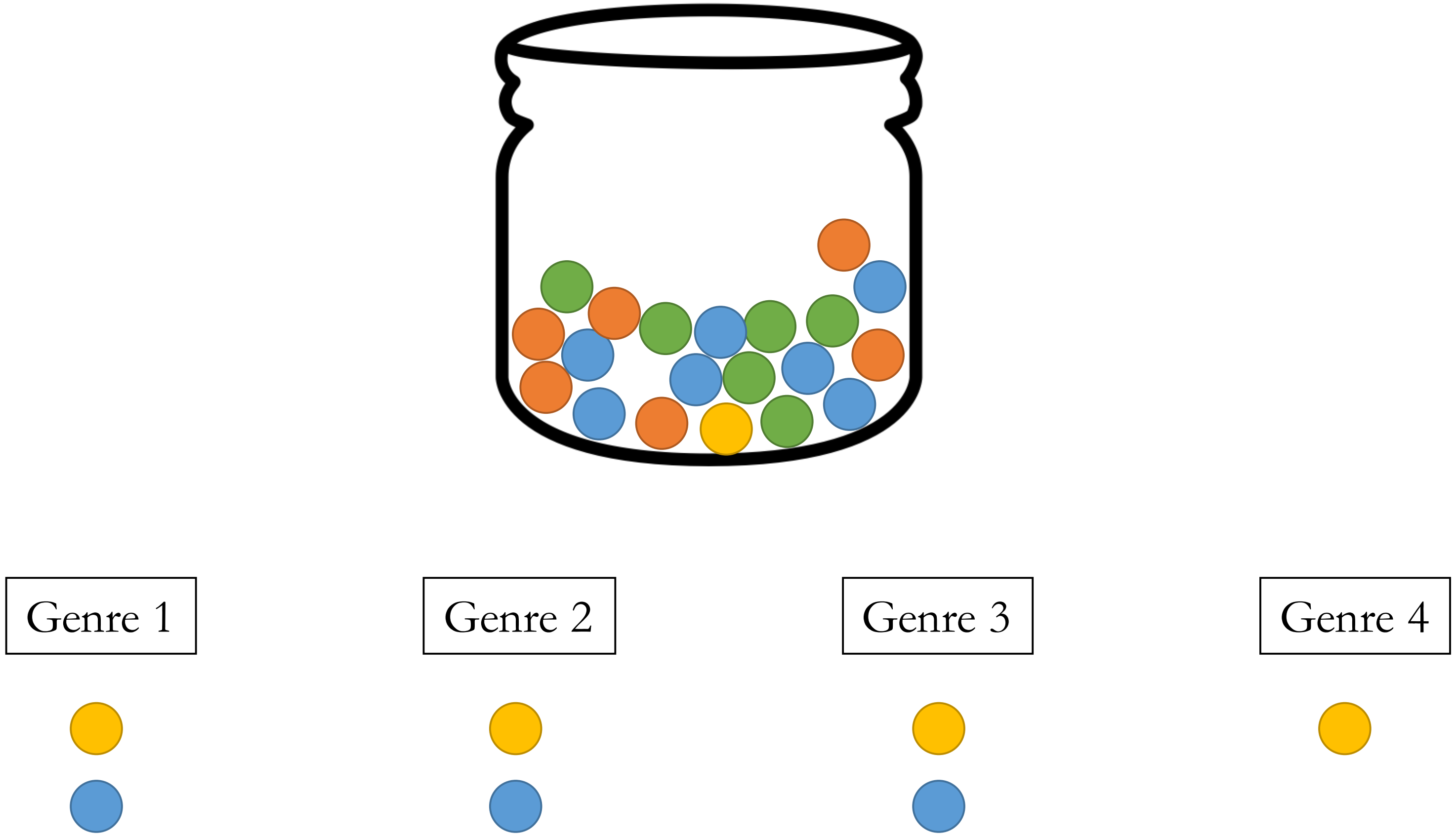
NYT Experiment: Matching



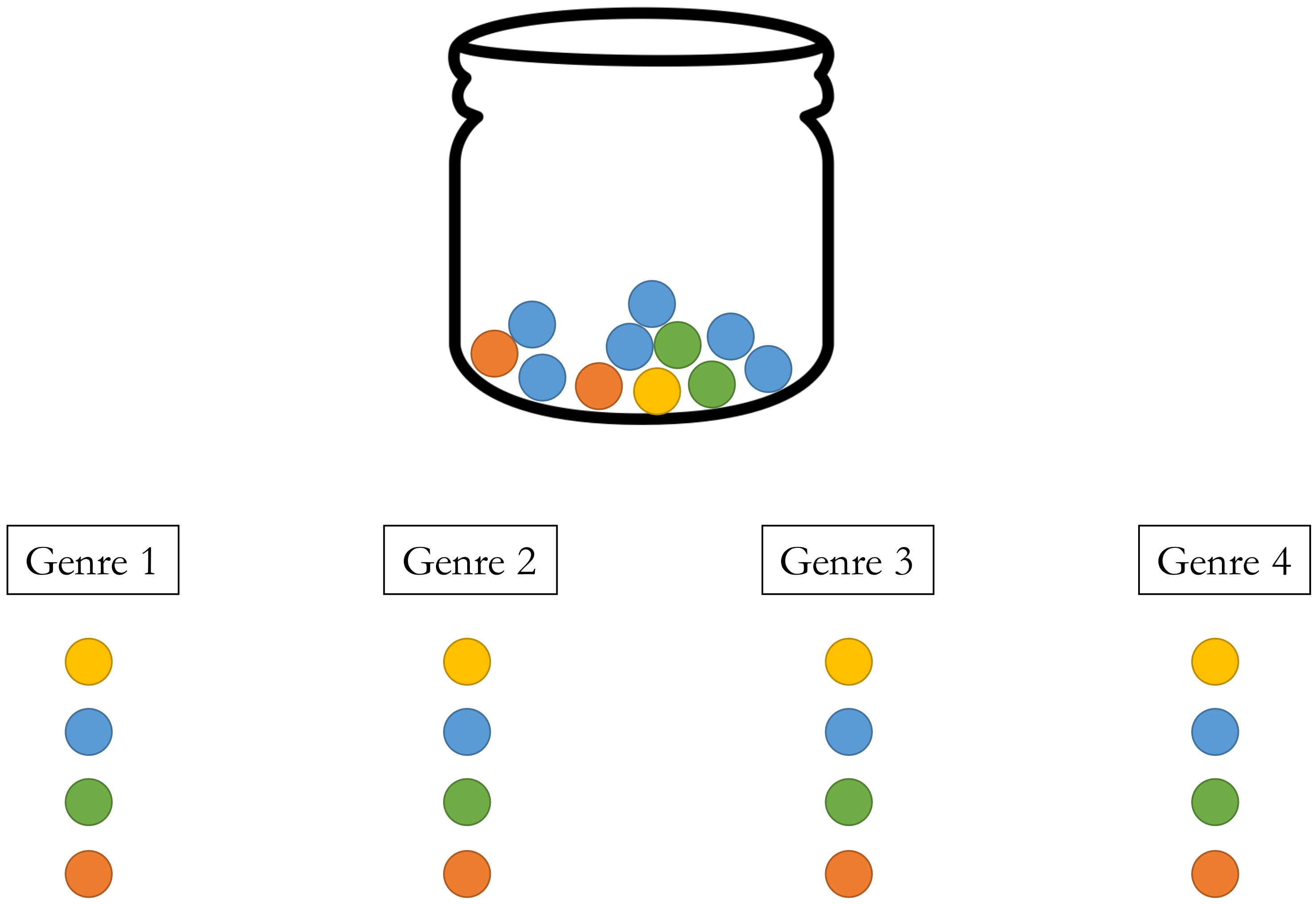
NYT Experiment: Matching



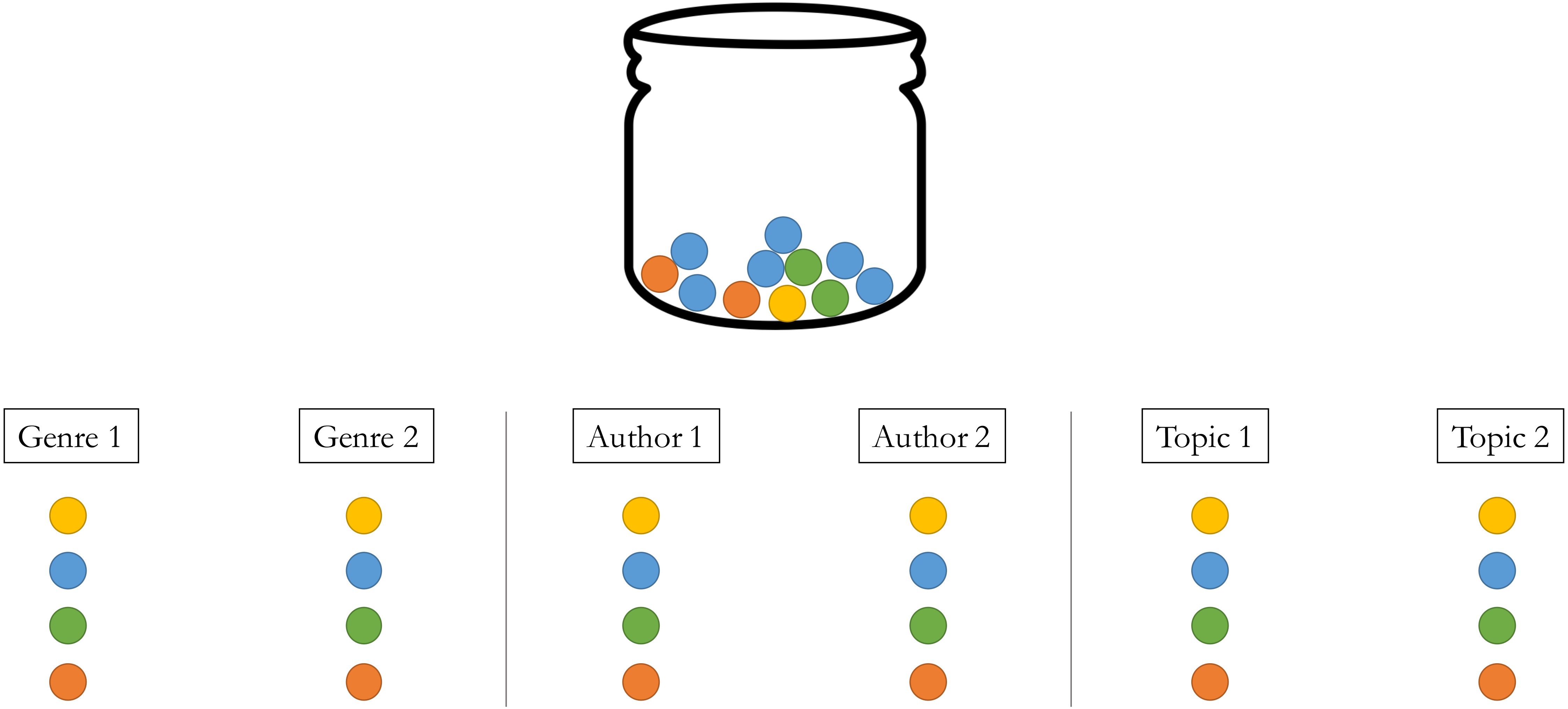
NYT Experiment: Matching



NYT Experiment: Matching



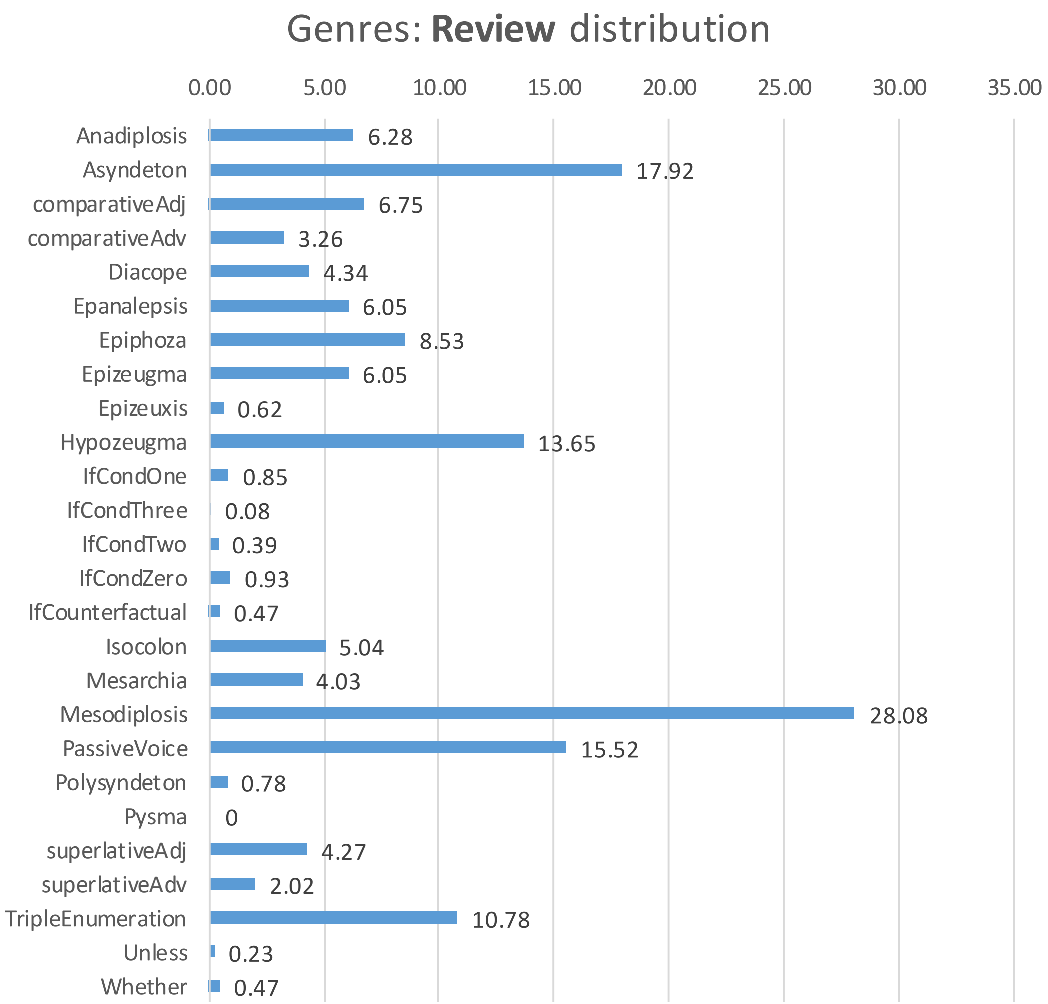
NYT Experiment: Matching

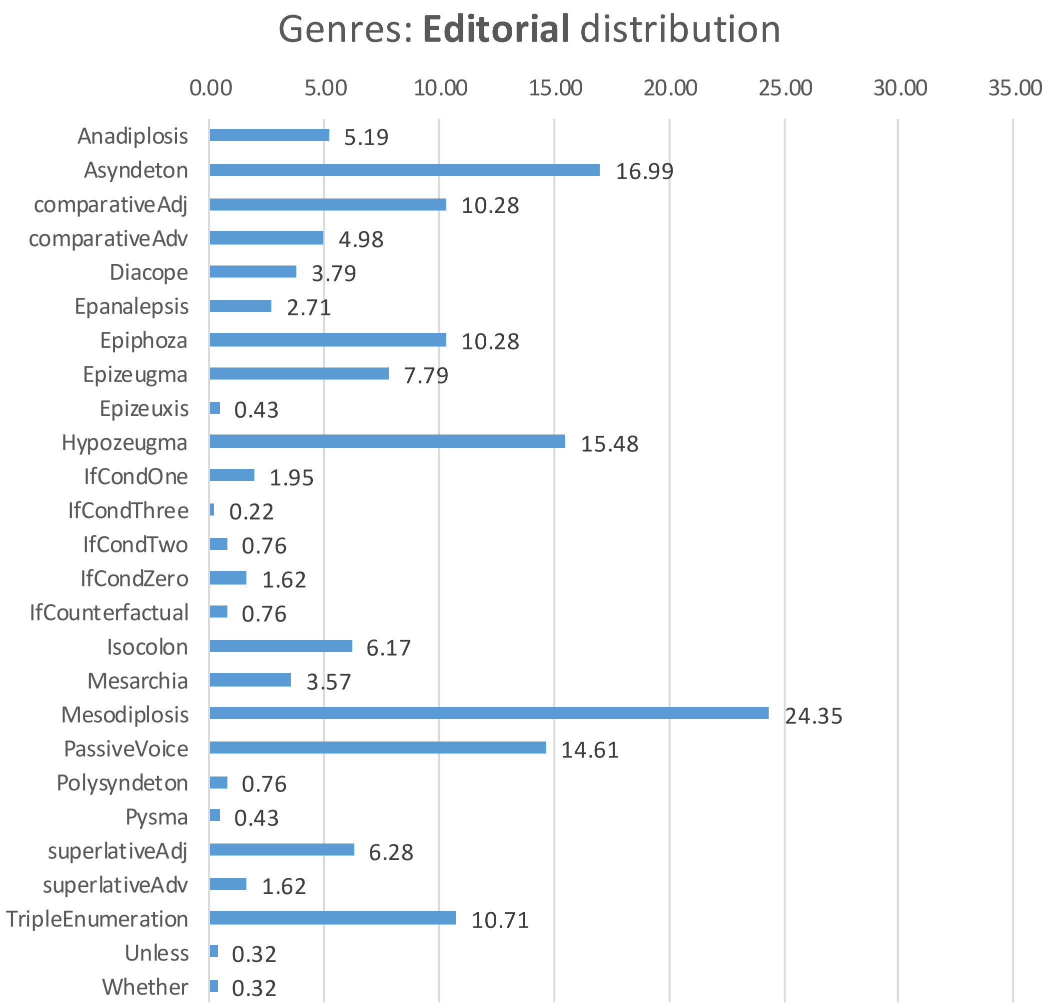


3 Analysis Experiments

Findings

NYT Experiment: Frequency





NYT Experiment: Findings

Style-based frequency of rhetorical devices

NYT Experiment: Findings

Authors

	EPIPHOZA	REPETITION SCHEMES
Author	**Distribution (%)**	**Distribution (%)**
Hevesi Dennis	10.74	70.99
Lewis Paul	12.99	81.93
Martin Douglas	6.49	55.49

NYT Experiment: Findings

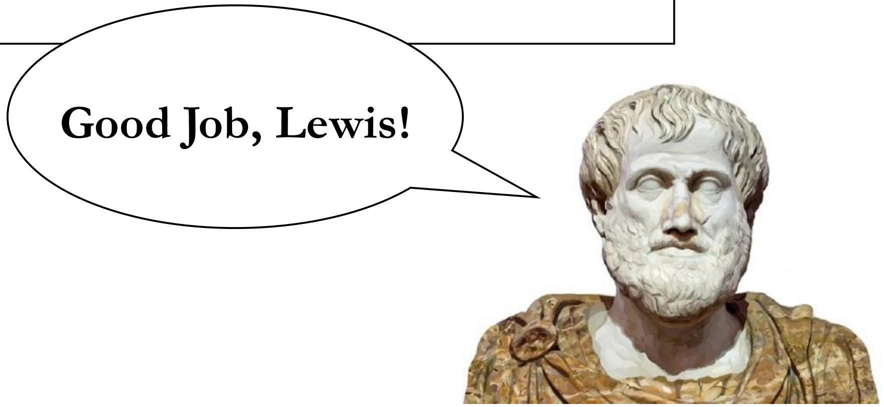
Authors

	EPIPHOZA	REPETITION SCHEMES
Author	**Distribution (%)**	**Distribution (%)**
Hevesi Dennis	10.74	70.99
Lewis Paul	12.99	81.93
Martin Douglas	6.49	55.49

! Same pattern across all articles

NYT Experiment: Findings

	Authors	
	EPIPHOZA	REPETITION SCHEMES
Author	**Distribution (%)**	**Distribution (%)**
Hevesi Dennis	10.74	70.99
Lewis Paul	12.99	81.93
Martin Douglas	6.49	55.49



NYT Experiment: Findings

Authors

	SIGNIFICANCE		EFFECT-SIZE	
Datasets	**P-value**	**Independence**	**Cramer's V value**	**Effect**
Hevesi vs. Lewis	0.015	TRUE*	0.1	SMALL
Lewis vs. Martin	$\sim$0	TRUE	0.15	SMALL
Martin vs. Hevesi	0.017	TRUE*	0.1	SMALL

* for $\alpha > 0.001$

NYT Experiment: Findings

Genres

Comparatives

Confounders	**Distribution (%)**		
Genre:	*Biography*	*Editorial*	*Review*
freedman-news	11.65	25.57	11.75
norris-markets	22.59	30.06	20.99
wade-health	12.04	12.97	16.40

author topic

NYT Experiment: Findings

Genres

	COMPARATIVES	CONDITIONALS
Genre	**Distribution (%)**	**Distribution (%)**
Biography	14.07	3.45
Editorial	23.16	5.95
Review	16.29	3.41

NYT Experiment: Findings

Genres: tests' results

	SIGNIFICANCE		EFFECT-SIZE	
Datasets	**P-value**	**Independence**	**Cramer's V value**	**Effect**
Biography vs. Editorial	∼0	TRUE	0.16	SMALL
Editorial vs. Review	∼0	TRUE	0.14	SMALL
Review vs. Biography	0.68	FALSE	0.07	SMALL

NYT Experiment: Findings

Topics

	COMPARATIVES		
Confounders	**Distribution (%)**		
Topics:	*Arts*	*Education*	*Science*
martin-biography	11.95	10.94	12.24
saxon-biography	6.15	6.14	12.50

author genre

NYT Experiment: Findings

Style-based frequency of rhetorical devices

Characteristic style patterns within each dimension

NYT Experiment: Findings

Topics: tests' results

	SIGNIFICANCE		EFFECT-SIZE	
Datasets	**P-value**	**Independence**	**Cramer's V value**	**Effect**
Science vs. Education	0.70	FALSE	0.09	SMALL
Education vs. Arts	0.26	FALSE	0.10	SMALL
Arts vs. Science	0.19	FALSE	0.10	SMALL

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

Comparatives

Debate Type	Distribution (%)
Clinton $\rightarrow$ Trump	11.00
Trump $\rightarrow$ Clinton	7.02

Presidential Debates: Findings

	ASYNDETON	VOICE	BALANCE SCH.
Debate Type	**Distribution (%)**	**Distribution (%)**	**Distribution (%)**
Clinton → Trump	15.24	8.07	17.69
Trump → Clinton	10.83	5.29	19.92

Presidential Debates: Findings

	ASYNDETON	VOICE	BALANCE SCH.
Debate Type	**Distribution (%)**	**Distribution (%)**	**Distribution (%)**
Clinton → Trump	15.24	8.07	17.69
Trump → Clinton	10.83	5.29	19.92

Asyndeton = clarity and rhythm

Presidential Debates: Findings

	ASYNDETON	VOICE	BALANCE SCH.
Debate Type	**Distribution (%)**	**Distribution (%)**	**Distribution (%)**
Clinton → Trump	15.24	8.07	17.69
Trump → Clinton	10.83	5.29	19.92

Acceptance Speech Analysis by Huffington Post

Candidate	Sent.	Long Sent. (%)	Passive voice (%)	Grade Level (US)
Hillary Clinton	413	7.26	3.39	5
Donald Trump	341	16.42	8.8	8

Presidential Debates: Findings

	ASYNDETON	VOICE	BALANCE SCH.
Debate Type	**Distribution (%)**	**Distribution (%)**	**Distribution (%)**
Clinton → Trump	15.24	8.07	17.69
Trump → Clinton	10.83	5.29	19.92

Acceptance Speech Analysis by Huffington Post

Candidate	**Sent.**	**Long Sent. (%)**	**Passive voice (%)**	**Grade Level (US)**
Hillary Clinton	413	7.26	3.39	5
Donald Trump	341	16.42	8.8	8

Presidential Debates: Findings

	ASYNDETON	VOICE	BALANCE SCH.
Debate Type	**Distribution (%)**	**Distribution (%)**	**Distribution (%)**
Clinton → Trump	15.24	8.07	17.69
Trump → Clinton	10.83	5.29	19.92

Acceptance Speech Analysis by Huffington Post

Candidate	Sent.	Long Sent. (%)	Passive voice (%)	Grade Level (US)
Hillary Clinton	413	7.26	3.39	5
Donald Trump	341	16.42	8.8	8

Presidential Debates: Findings

Significance Test

Debate Type	*Clinton → Rest*	*Clinton → Trump*	*Trump → Clinton*	*Trump → Rest*
Clinton → Rest		TRUE*	TRUE	TRUE
Clinton → Trump	TRUE*		TRUE	TRUE
Trump → Clinton	TRUE	TRUE		FALSE†
Trump → Rest	TRUE	TRUE	FALSE†	

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

Presidential Debates: Findings

Significance Test

Debate Type	*Clinton → Rest*	*Clinton → Trump*	*Trump → Clinton*	*Trump → Rest*
Clinton → Rest		TRUE*	TRUE	TRUE
Clinton → Trump	TRUE*		TRUE	TRUE
Trump → Clinton	TRUE	TRUE		FALSE†
Trump → Rest	TRUE	TRUE	FALSE†	

* 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	Resources
System for rhetorical style identification in high-quality text documents	Novel framework for detecting rhetorical devices
Rule-based algorithms for detection of RD	Comprehensive dataset for evaluation of rhetoric detection systems
Vague style patterns across random and article-length based subsampling: **Confounding**	Elaborative style patterns and intriguing findings
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

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

- 1st sentence → 5.8 sec.
- 2nd sentence → 0.4 sec.
- Initialization → 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

- Ben Wiseman, New York Times https://www.nytimes.com/2016/09/25/opinion/campaign-stops/my-debate-nightmare-a-duller-donald-trump.html
- Peter Kluegl and Martin Atzmueller. Textmarker: A tool for rule-based infor-mation extraction, 2009.
- Khalid Al-Khatib, Henning Wachsmuth, Matthias Hagen, and Benno Stein. Patterns of Argumentation Strategies across Topics. In Proceedings of the 2017 Conference on Empirical Methods in Natural Language Processing (EMNLP 17), pages 1362–1368. Association for Computational Linguistics, September 2017. URL http://aclweb.org/anthology/D17-1142.
- 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
- Folder 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 Llisole 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
- http://community.wikia.com/wiki/File:Aristotle-17.jpg
- https://www.washingtonpost.com/graphics/politics/2016-election/presidential-debate-schedule/

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

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

- Balance schemes   - Omission schemes   - Repetition schemes   - Custom schemes

Evaluation results

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

- Balance schemes - Omission schemes - Repetition schemes - Custom schemes

Evaluation results

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

- Balance schemes - Omission schemes - Repetition schemes - Custom schemes

Evaluation results

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

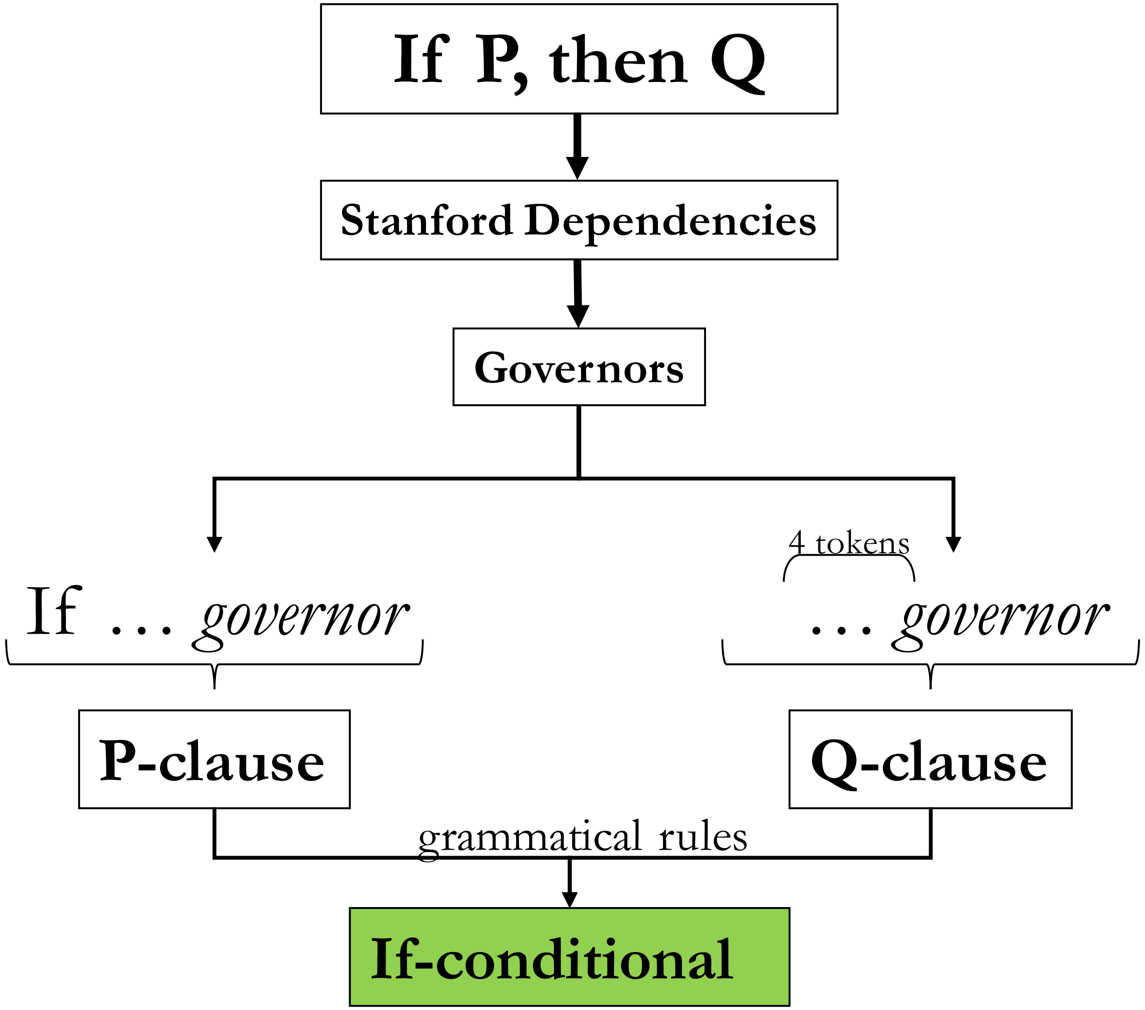
- Balance schemes - Omission schemes - Repetition schemes - Custom schemes

Evaluation results

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

- Balance schemes - Omission schemes - Repetition schemes - Custom schemes

If-conditional Detection



If-counterfactual Detection