Rhetorical Figure Detection: the Case of Chiasmus

MARIE DUBREMETZ & JOAKIM NIVRE



UPPSALA UNIVERSITY SWEDEN

Rhetoric workshop - February 2016

Stylometric tools: example



• Match string, words, shallow pattern

More? Rhetorical devices? Chiasmus?

Stylometric tools: example



- Match string, words, shallow pattern
- More? Rhetorical devices? Chiasmus?

Introduction

Chiasmus/Antimetabole: Traditional Definition

A rhetorical figure in which two words are repeated in reverse order.

Introduction

Chiasmus/Antimetabole: Traditional Definition

A rhetorical figure in which two words are repeated in reverse order.





- Help corpus linguistics and rhetorician projects.
- Improve our general knowledge of the figure?
- If we can make it for chiasmus, you can hope to make it for more devices.



• Help corpus linguistics and rhetorician projects.

- Improve our general knowledge of the figure?
- If we can make it for chiasmus, you can hope to make it for more devices.



- Help corpus linguistics and rhetorician projects.
- Improve our general knowledge of the figure?
- If we can make it for chiasmus, you can hope to make it for more devices.



- Help corpus linguistics and rhetorician projects.
- Improve our general knowledge of the figure?
- If we can make it for chiasmus, you can hope to make it for more devices.

Question



Why does the chiasmus detector not exist yet?

- Lack of interest?
- Impossible technically?

Question



Why does the chiasmus detector not exist yet?

Lack of interest?

Impossible technically?

Question



Why does the chiasmus detector not exist yet?

- Lack of interest?
- Impossible technically?

Problem

Traditional definition

"Chiasmus: reversal of two pairs of words"

This is a definition of necessary but not sufficient conditions

In other words

Chiasmus is not a trivial linguistic device. Can a computer detect it?

Problem

Traditional definition

"Chiasmus: reversal of two pairs of words"

• This is a definition of necessary but not sufficient conditions

In other words

Chiasmus is not a trivial linguistic device. Can a computer detect it?

The research on chiasmus

 Gawryjolek [2009]: Extract every double pair of words with reverse order without exception Chuck Norris does not fear death, death fears Chuck Norris

- 100% recall
- Very low precision (< 1%)
- Hromada [2011]: Identify not two but three pairs of reverted words

Love makes time pass, time makes love pass.

- Very high precision
- But low recall

Problem

There are criss-cross patterns that are not chiasmi such as:

'I like beer from time to time but I prefer wine'

They are frequent but chiasmi are rare.

Problem

There are criss-cross patterns that are not chiasmi such as:

'I like beer from time to time but I prefer wine'

They are frequent but chiasmi are rare.

Problem

There are criss-cross patterns that are not chiasmi such as:

'I like beer from time to time but I prefer wine'

They are frequent but chiasmi are rare.

Problem

There are criss-cross patterns that are not chiasmi such as:

'I like beer from time to time but I prefer wine'

They are frequent but chiasmi are rare.



This is the problem of the needle in the haystack!

(Re-)Defining the Task

Both pair [2009] and triplet [2011] solutions are extreme...



(Re-)Defining the Task

Both pair [2009] and triplet [2011] solutions are extreme...



... Why not outputting chiasmi in a sorted manner?

Thus the user can choose to read only a couple of pages of results (instead of thousands!)

Marie Dubremetz & Joakim Nivre

How Do We Sort?

Any ranking is based on a score.



How does the computer score a chiasmus?

How Do We Rank?

We split the judgement into micro evaluations...



How Do We Rank?

We split the judgement into micro evaluations...



Thus the task becomes computer friendly.

How Do We score? An Example of Features

First should be last and

last should be first.

I like beer from time to

time . But I prefer wine.

How our algorithm sorts criss-cross patterns: 5 representative examples of our 20 features

How Do We Score? An Example of Features



How our algorithm sorts criss-cross patterns: 5 representative examples of our 20 features

How Do We Score? An Example of Features



How our algorithm sorts criss-cross patterns: 5 representative examples of our 20 features

Our Model

A standard linear model

We encode 4 different types of features:

- Basic (stopwords and punctuation)
- Size
- Ngram
- Lexical Features

We give a score to every criss-cross pattern relative to those features.

This score allows us to rank chiasmi.

To evaluate the system we asked a human user to annotate

• 800 criss-cross patterns

To evaluate the system we asked a human user to annotate

- 800 criss-cross patterns
- True: Food should be our medicine and medicine our food.

To evaluate the system we asked a human user to annotate

- 800 criss-cross patterns
- True: Food should be our medicine and medicine our food.
- *Borderline*: If all this is not **helped** by a **fund**, the **fund** is no **help** at all.

To evaluate the system we asked a human user to annotate

- 800 criss-cross patterns
- True: Food should be our medicine and medicine our food.
- *Borderline*: If all this is not **helped** by a **fund**, the **fund** is no **help** at all.
- *False*: A ban is **already** in place in **several** Member States, as **several** speakers have **already** mentioned.

- No ranking system in previous work.
- 9 candidates only in [2011]
- Thus table limited to precision at 9.





- No ranking system in previous work.
- 9 candidates only in [2011]
- Thus table limited to precision at 9.

Comparison table			
	Our	3	2
	system	pairs	pairs
	[2015]	[2011]	[2009]
Precision	7/9	6/9	0/9
Precision (%)	78	67	0
	•		



- No ranking system in previous work.
- 9 candidates only in [2011]
- Thus table limited to precision at 9.



- => Precision comparable to [2011]
- => And improved recall (17/(+-)22)

Future Work

- What about syntactic features?
- More annotation with more annotators
- Apply our method to other devices? Anaphora? Anadiplosis?

 A new approach to a rhetorical figure: chiasmus as a graded phenomenon

- 2 Pushed the theoretical limits of computers
- More work for rhetoricians and linguists! Your judgement is required

and

- Bible, 800 000 words
- Sherlock Holmes, 650 000 words

- A new approach to a rhetorical figure: chiasmus as a graded phenomenon
- 2 Pushed the theoretical limits of computers
- More work for rhetoricians and linguists! Your judgement is required

and

- Bible, 800 000 words
- Sherlock Holmes, 650 000 words

- A new approach to a rhetorical figure: chiasmus as a graded phenomenon
- 2 Pushed the theoretical limits of computers
- More work for rhetoricians and linguists! Your judgement is required

and

- Bible, 800 000 words
- Sherlock Holmes, 650 000 words

- A new approach to a rhetorical figure: chiasmus as a graded phenomenon
- 2 Pushed the theoretical limits of computers
- More work for rhetoricians and linguists! Your judgement is required

and

- Bible, 800 000 words
- Sherlock Holmes, 650 000 words

Thank You!

Questions?

All for one, one for all. (In binary.)

Gawryjolek, J. J. (2009). Automated Annotation and Visualization of Rhetorical Figures. Master thesis, University of Waterloo.

Hromada, D. D. (2011). Initial Experiments with Multilingual Extraction of Rhetoric Figures by means of PERL-compatible Regular Expressions. In *Proceedings of the Second Student Research Workshop associated with RANLP 2011*, (pp. 85–90)., Hissar, Bulgaria.

System, corpus, annotation available at:

http://stp.lingfil.uu.se/~marie/chiasme.htm.

References

System, corpus, annotation available at: http://stp.lingfil.uu.se/~marie/chiasme.htm.