### The Translators' Touch

A Computational Stylometric Inquiry into

Medieval Greek-Latin Translations

Pieter Beullens (FWO Vlaanderen - KU Leuven) Wouter Haverals (UAntwerpen) Ben Nagy (Polish Academy of Sciences)

#### Research Question:

Can computational stylometry help to identify translators, and corroborate or contest previous attributions?

#### Research Question: Challenge

How to distinguish the Latin translators' touch from the style of the Greek authors?

# Background & History

#### Corpus

- Translators from Greek into Latin from the 12th and 13th centuries
- Mainly Aristotle, but also other texts on philosophy, theology, medicine, and astrology

artyfere pa outfum of the are in the property of the continues of the party of the continues of the party of the continues of

crecini qui of tomic of control of control of the control of this control of this control of the control of the

ruit he am & comme Couly ale far gor renew was inf pur and hold te acceptioning prime a anni prot another fito hamb' priagonoi priago notwork the and cog Progt flatoni to hea'co mrea le gamne Gamp' forma a'ona oligazona fiir excepta con three Halling pie gaplotene vila ed og tatuna prontonna puna anna quidmino ga inuolacilia no imois pu enpoure remier Amarara mana frir n's nond'tur Bamd'of facuir mortien อักร และอยุมสิตา ผู้สามอยุทา เนตกอนอ go mforamm mein tofflif montreat br Mamera mound facuit derried 1 net crimos inpalvartoz co fectaria a un mina thicof che podemicare h E denound my egojs a four n & E mon morticati se hm govens thepe almoneauf warmmet deverans hirmisaramorio monficar time" ceri a'elt qiinucamo epenilernica mantia morficar mmarous finistaris com

min a omili multi omin felindina
off hymicuthers up activité of
ann amplita publishine souvered
a vinamin es auto thi int 'n' quot un
teil Distinut a fait tuins of summer
stragus apen a montain frecomment
the auto is anitation women theta a
thu souver a consensition to omit
the auto is a mitation women the and
it has guid in alto souver of a faitha
a tout of account pun a unitation
thas a vicion pun a unitation
this in a minerit tuin his e
teil ship a number tuin his e
teil ship a tuin souver
this noul selicia in e d
in a cut it is in minerit teautte a faout and lass like a language
tuin souver a la cut a la
tuin del ship a minerit tuin ship
tuin del ship
tuin

#### Dramatis Personae

#### 12th century

- James of Venice
- Burgundio of Pisa
- Henry Aristippus
- Anonymous translator (Physica Vaticana and Metaphysica media)

#### 13th century

- Bartholomew of Messina
- William of Moerbeke

#### Evidence for identification

- Some translations are attributed in manuscripts (trustworthy?)
- Some translations are transmitted together
- Identification through stylistic analysis

#### Proto-history

- Long-standing tradition of stylometric investigations
  - Lorenzo Minio-Paluello on James of Venice and Henry Aristippus
  - Fernand Bossier on William of Moerbeke & Burgundio of Pisa
- Content words: might be influenced by the cultural environment or vary during the career of a translator

#### Proto-history

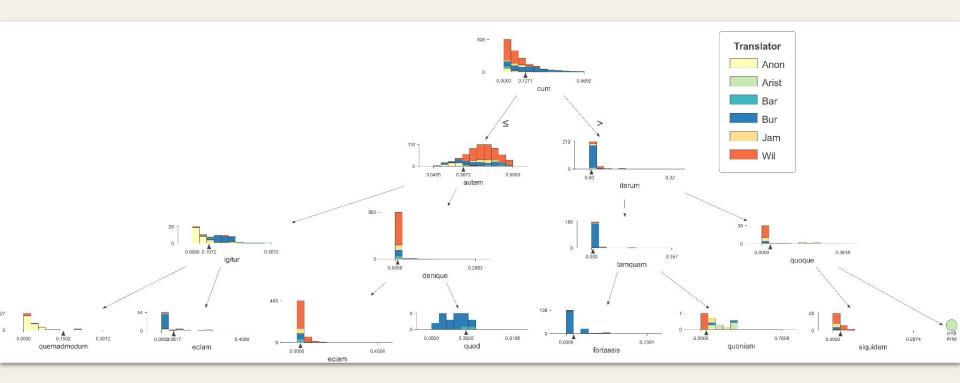
- Function words:
  - Very strict translation method: word-for-word
  - Greek language uses many modifiers
  - Latin translations deviate from classical Latin idiom to mirror the Greek source texts
- Research based on equivalence between Greek and Latin & frequencies

#### Research Question: Refined

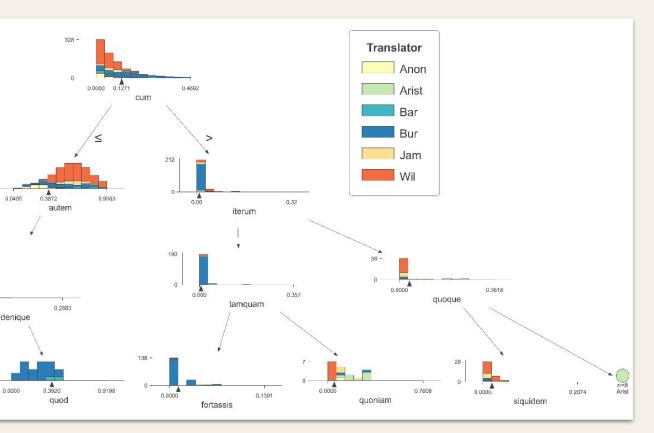
Which features are most discriminatory for translators, independent from the source texts?

## General Findings

#### Decision Tree Classifiers—How Do They Work?



#### Decision Tree Classifiers—How Do They Work?





## One bad apple...

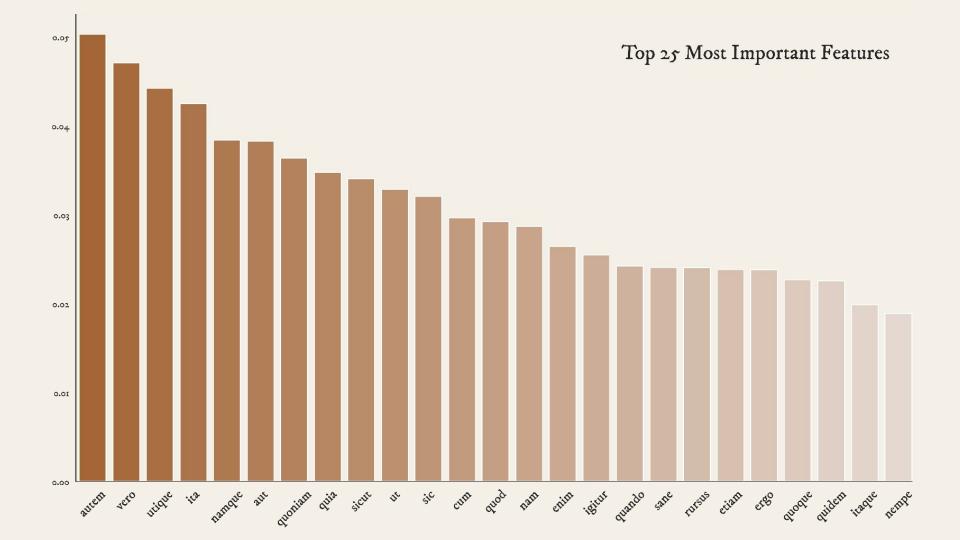
#### Problems with unbalanced corpora

	precision	recall	fi-score	support		precision	recall	f1-score	support
Anon	1.00	1.00	1.00	11	Anon	1.00	1.00	1.00	11
Arist	1.00	1.00	1.00	7	Arist	1.00	1.00	1.00	6
Bar	1.00	0.62	0.77	8	Bar	1.00	1.00	1.00	8
Bur	0.99	1.00	0.99	75	Bur	1.00	1.00	1.00	76
Jam	1.00	0.75	0.86	16	Jam	1.00	0.81	0.90	16
Wil	0.95	1.00	0.97	115	Wil	0.97	1.00	0.99	115
accuracy	0.97	0.97	0.97	0	accuracy	0.99	0.99	0.99	0
macro avg	0.99	0.90	0.93	232	macro avg	1.00	0.97	0.98	232
weighted avg	0.97	0.97	0.97	232	weighted avg	0.99	0.99	0.99	232





# Function Word 'Keyness'



# A Stylometric Investigation

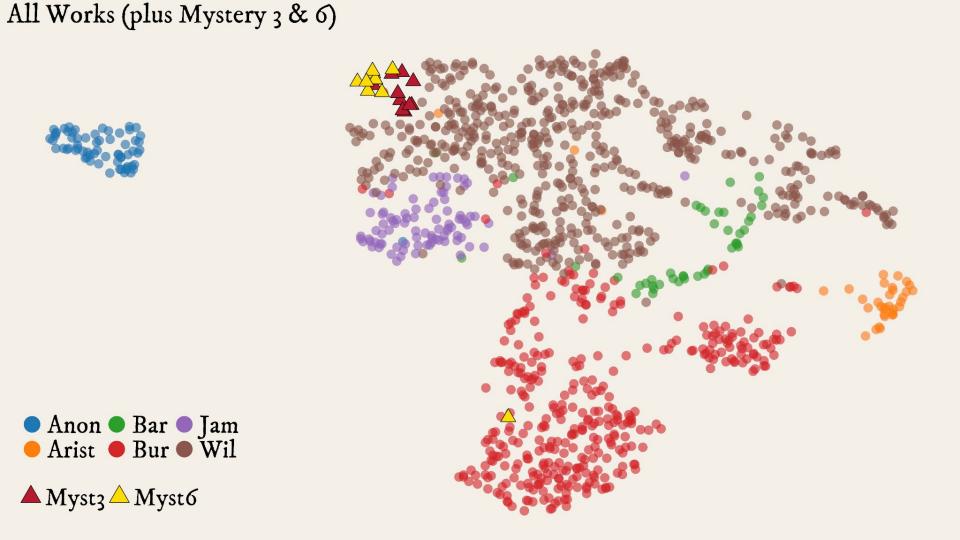
#### Investigation Background

- Translation of Sextus Empiricus Pros Mathematikous
- Some see a link with Bartholomew, some suggest an unknown party
- Comparison text of Sextus, probably by the same translator
- What does the computational evidence support?

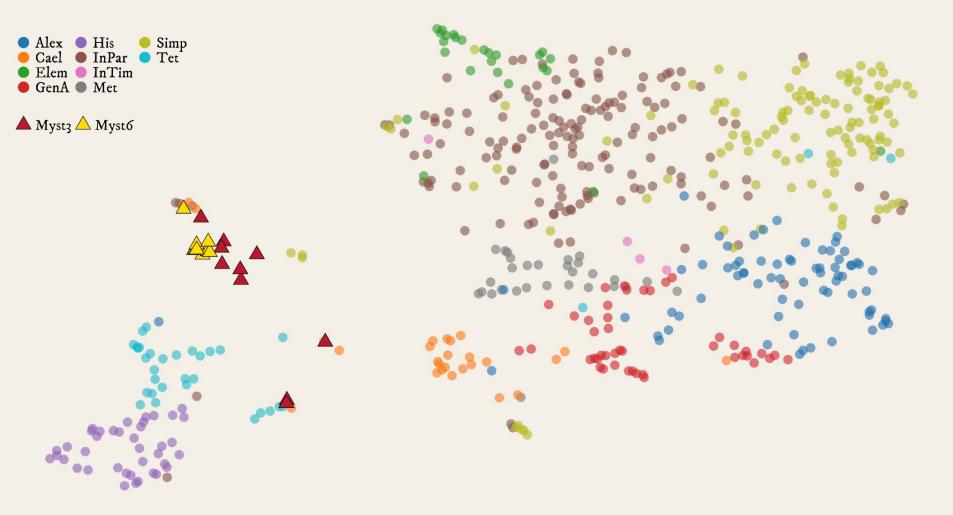
## Cluster Analysis

#### Methods

- Calculate function-word frequencies, L2 normalised
- Standard scaling (z-score) because frequencies vary a lot
- Project to 2 dimensions with UMAP



#### William of Moerbeke



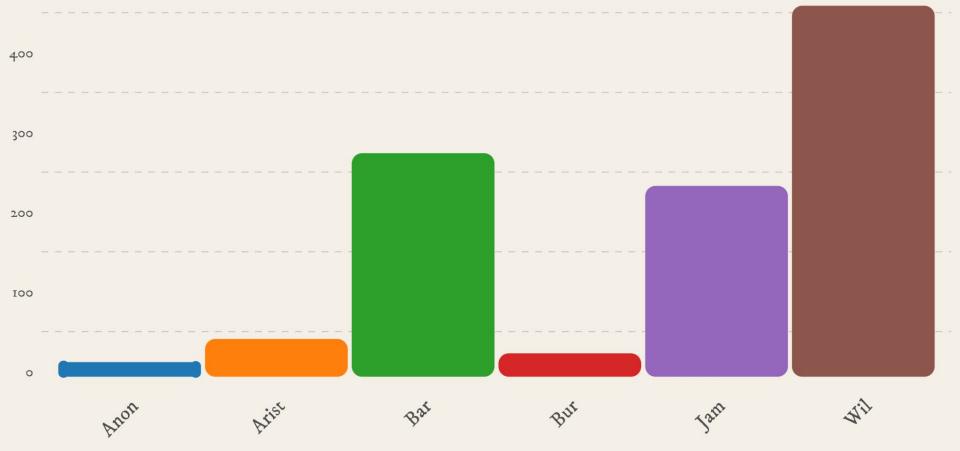
# Bartholomew of Messina HomMunSigMirabPri MorPue ▲ Myst3 △ Myst6

# Bootstrap Decision Trees

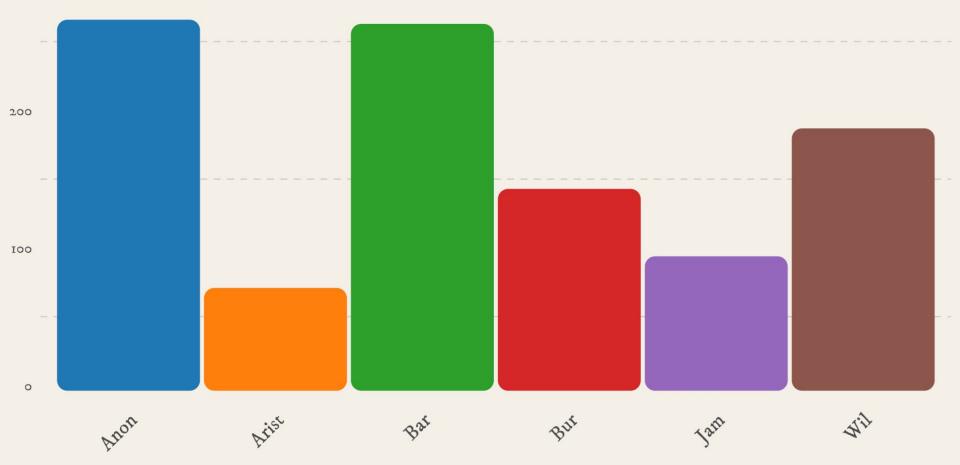
#### Methods

- 20 balanced corpus samples (17 x 1000-word chunks per translator)
- One Random Forest classifier per sample
- Classify 50 random subsamples from Myst3
- Count 'votes'

#### Bootstrap Classification of Myst3 (1000 votes)



#### Bootstrap Classification of Myst6 (1000 votes)

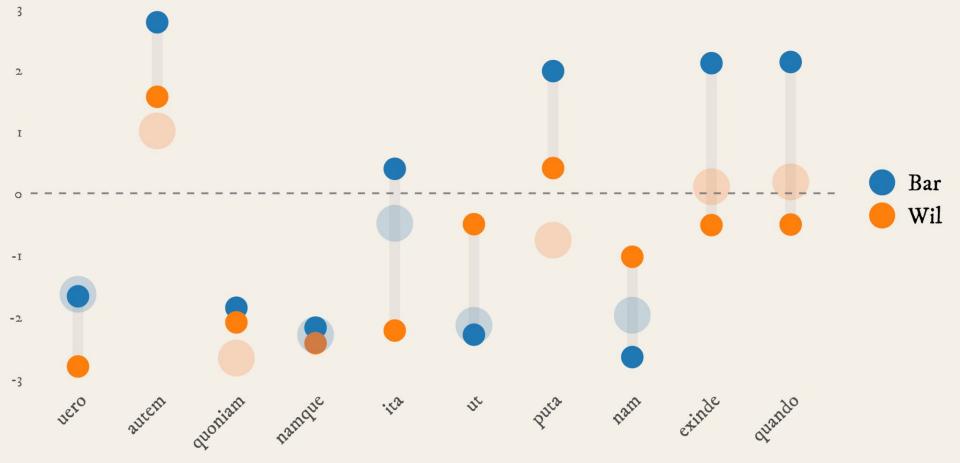


# Key Function Words

#### Methods

- Calculate Mutual Information for FWs, translator-by-translator
- Choose words with best combined MI
- NOTE: true joint MI is not as useful!

#### 10 key function-words (rescaled)



## Conclusions

#### A Secret, Third Thing?

#### Observatio

- Huge variation in classification of Myst3 and Myst6
- Myst 3 & 6 cluster together, but not closely with any translator

#### Argumentum

- IF they are by the same translator, unlikely it is one of our six
- Focused translator verification (not attribution) study required

# Discussion!