

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?

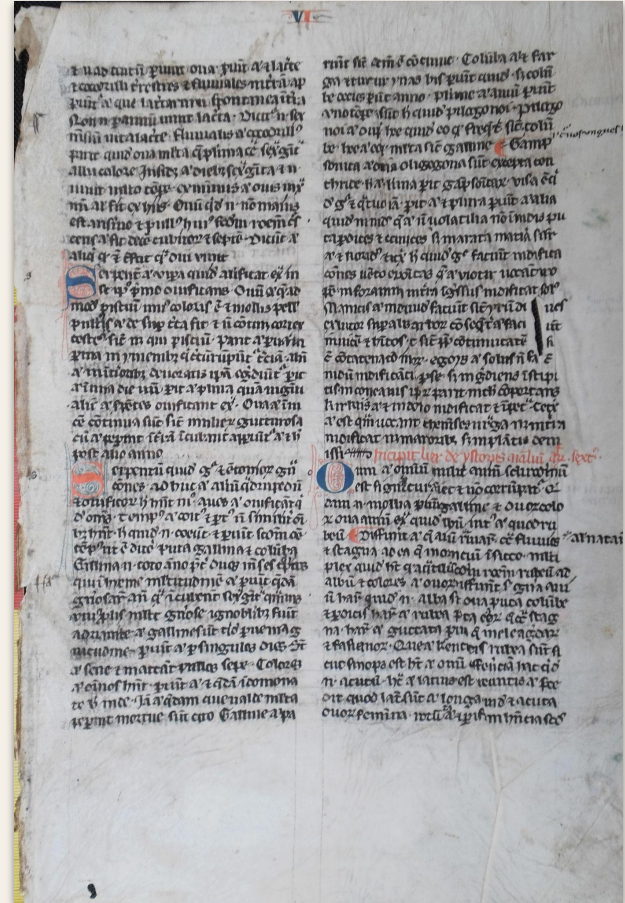
Research Question: Challenge

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



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

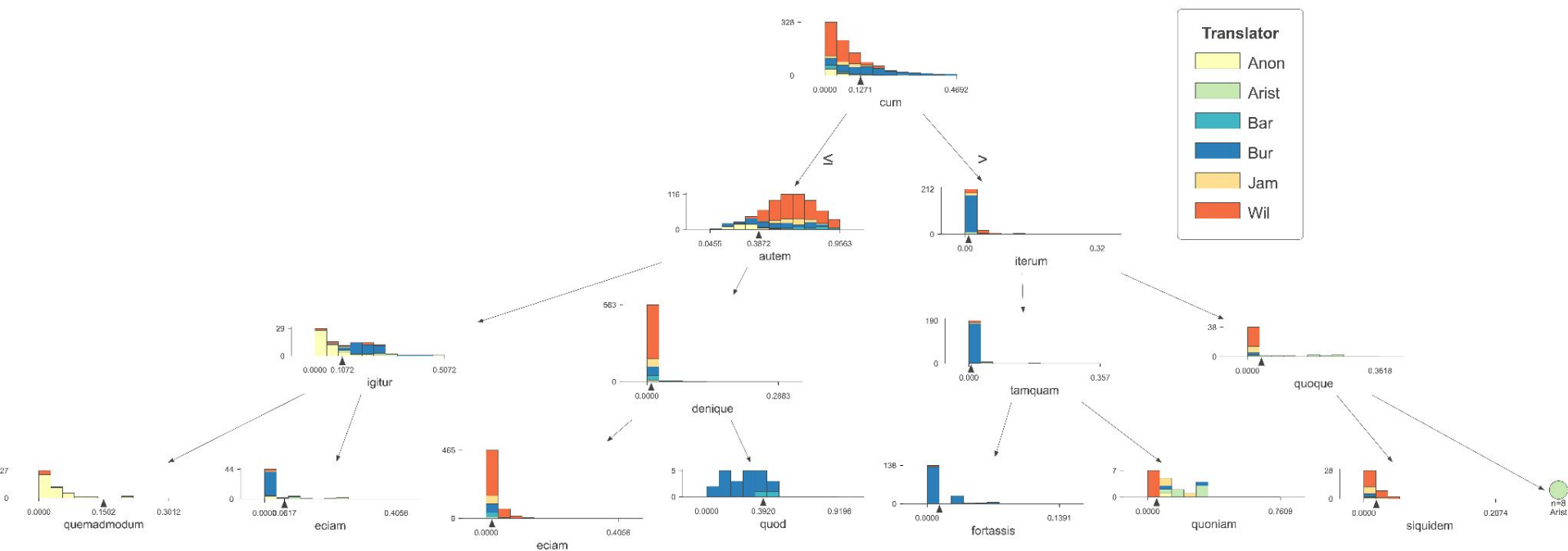
- Function words:
 - Very strict translation method: word-for-word
 - Greek language uses many modifiers
 - Latin translations deviate from classical Latin idiom to mirror the Greek source texts
- Research based on equivalence between Greek and Latin & frequencies

Research Question: Refined

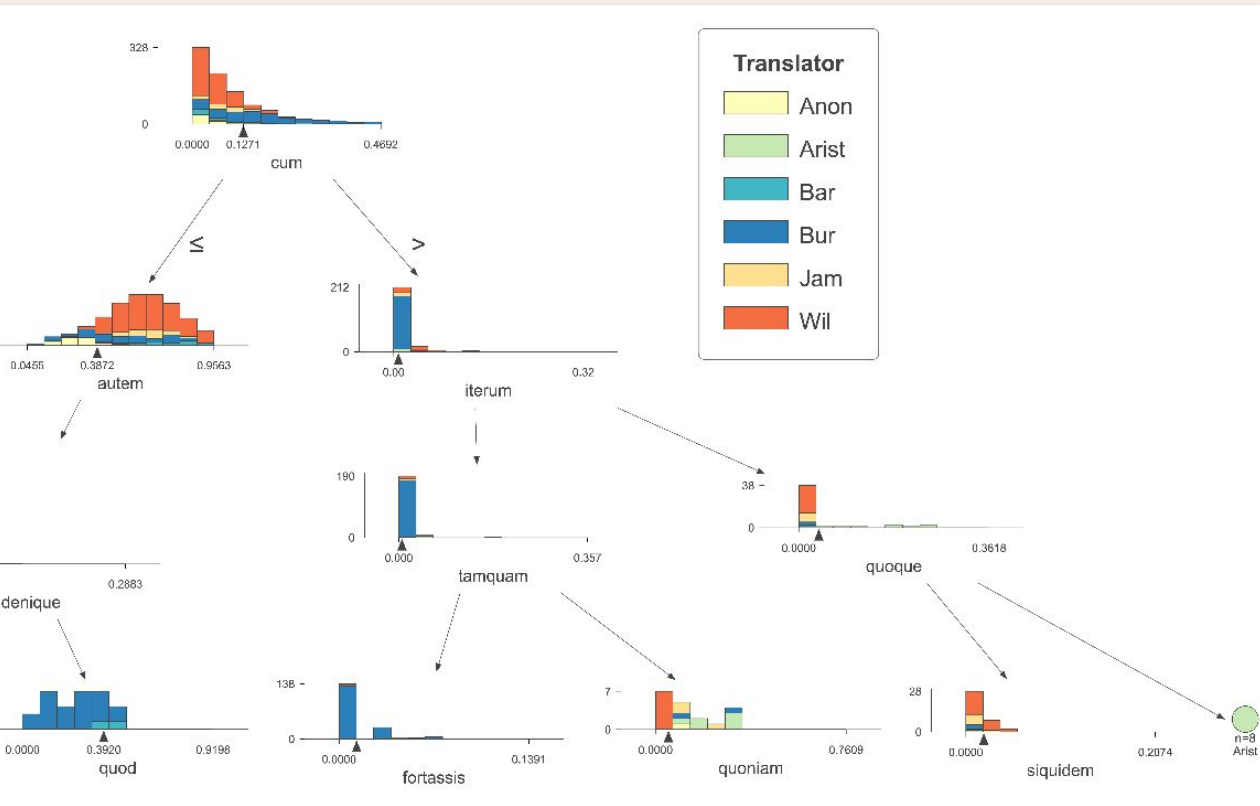
Which features are most discriminatory for translators, independent from the source texts?

General Findings

Decision Tree Classifiers—How Do They Work?



Decision Tree Classifiers—How Do They Work?

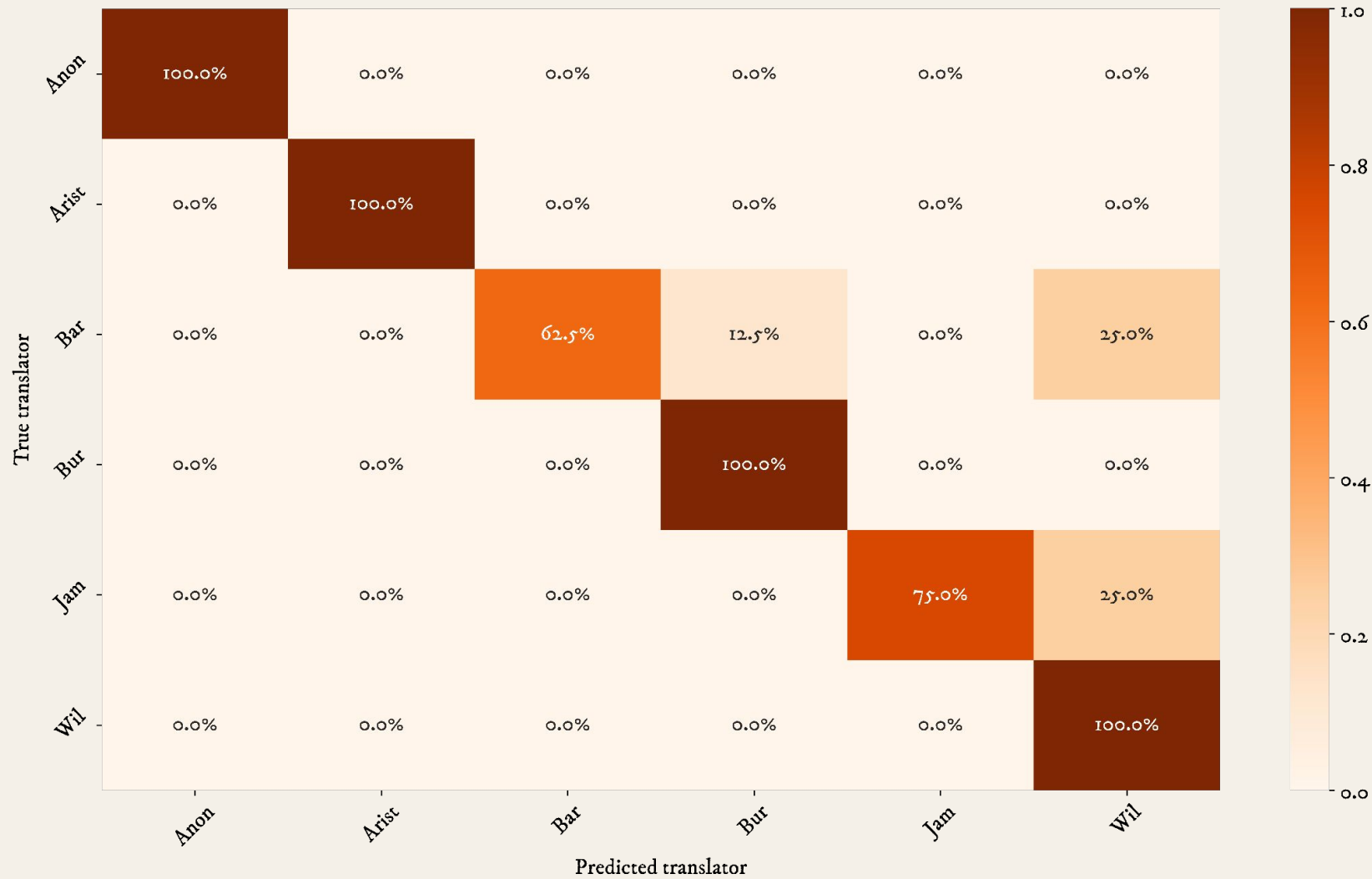


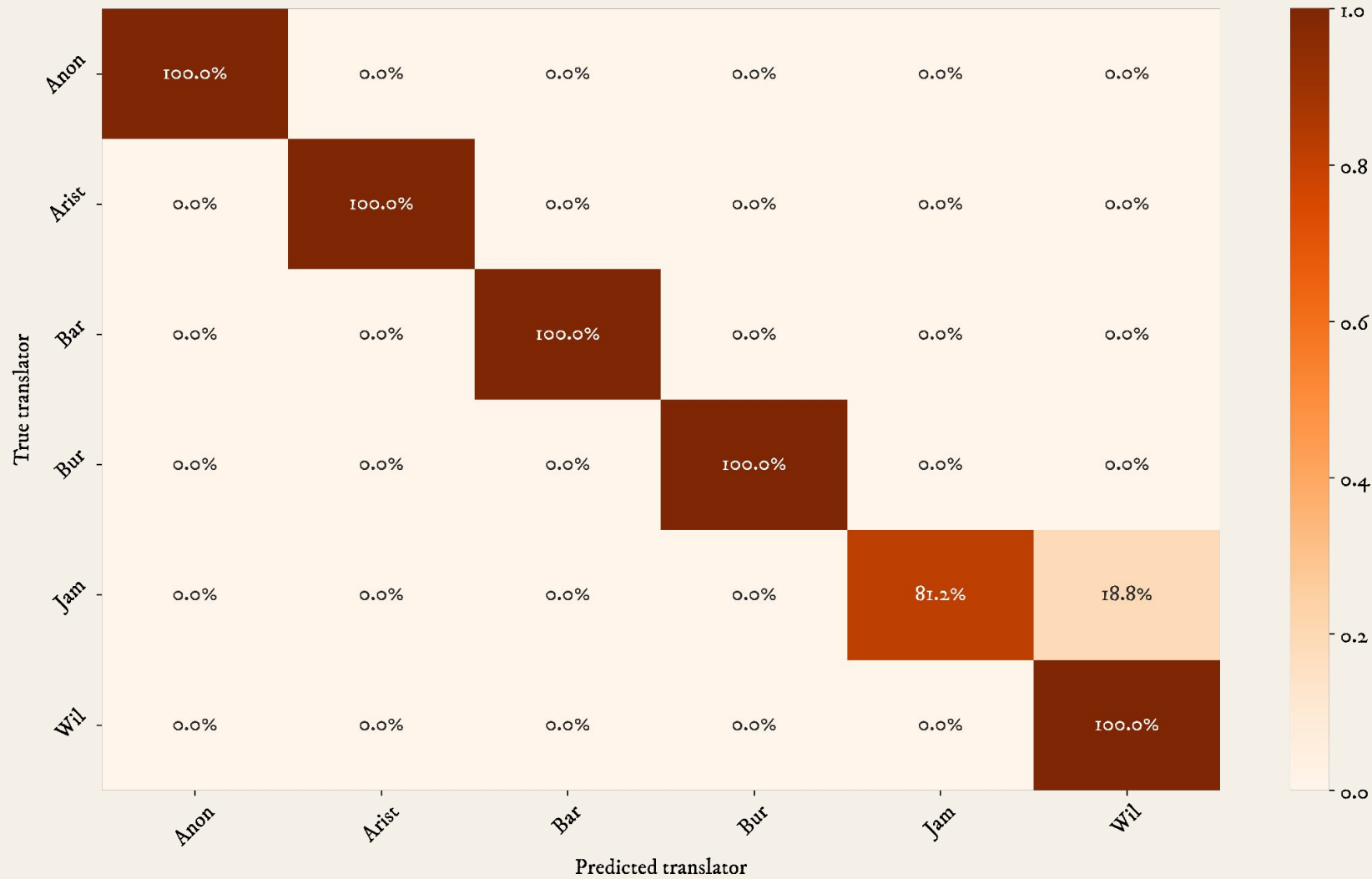
One bad apple...

Problems with unbalanced corpora

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

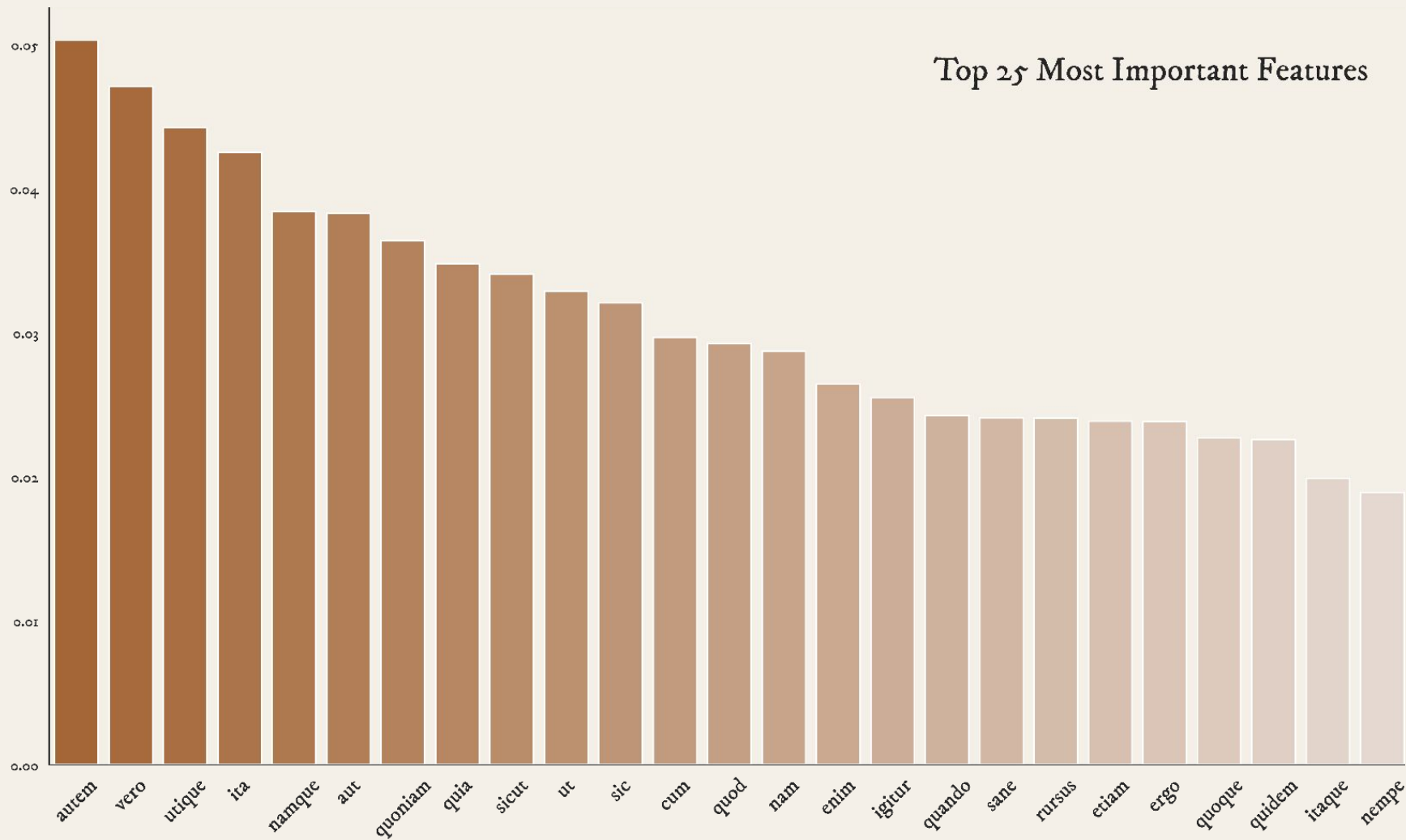
	precision	recall	f1-score	support
Anon	1.00	1.00	1.00	11
Arist	1.00	1.00	1.00	6
Bar	1.00	1.00	1.00	8
Bur	1.00	1.00	1.00	76
Jam	1.00	0.81	0.90	16
Wil	0.97	1.00	0.99	115
accuracy	0.99	0.99	0.99	0
macro avg	1.00	0.97	0.98	232
weighted avg	0.99	0.99	0.99	232





Function Word
'Keyness'

Top 25 Most Important Features



A Stylometric Investigation

Investigation Background

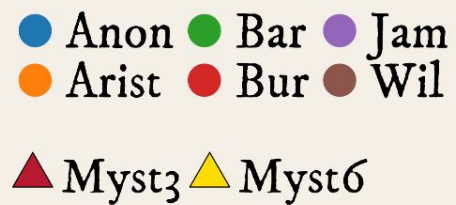
- Translation of Sextus Empiricus *Pros Mathematikous*
- Some see a link with Bartholomew, some suggest an unknown party
- Comparison text of Sextus, probably by the same translator
- What does the computational evidence support?

Cluster Analysis

Methods

- Calculate function-word frequencies, L_2 normalised
- Standard scaling (z -score) because frequencies vary a lot
- Project to 2 dimensions with UMAP

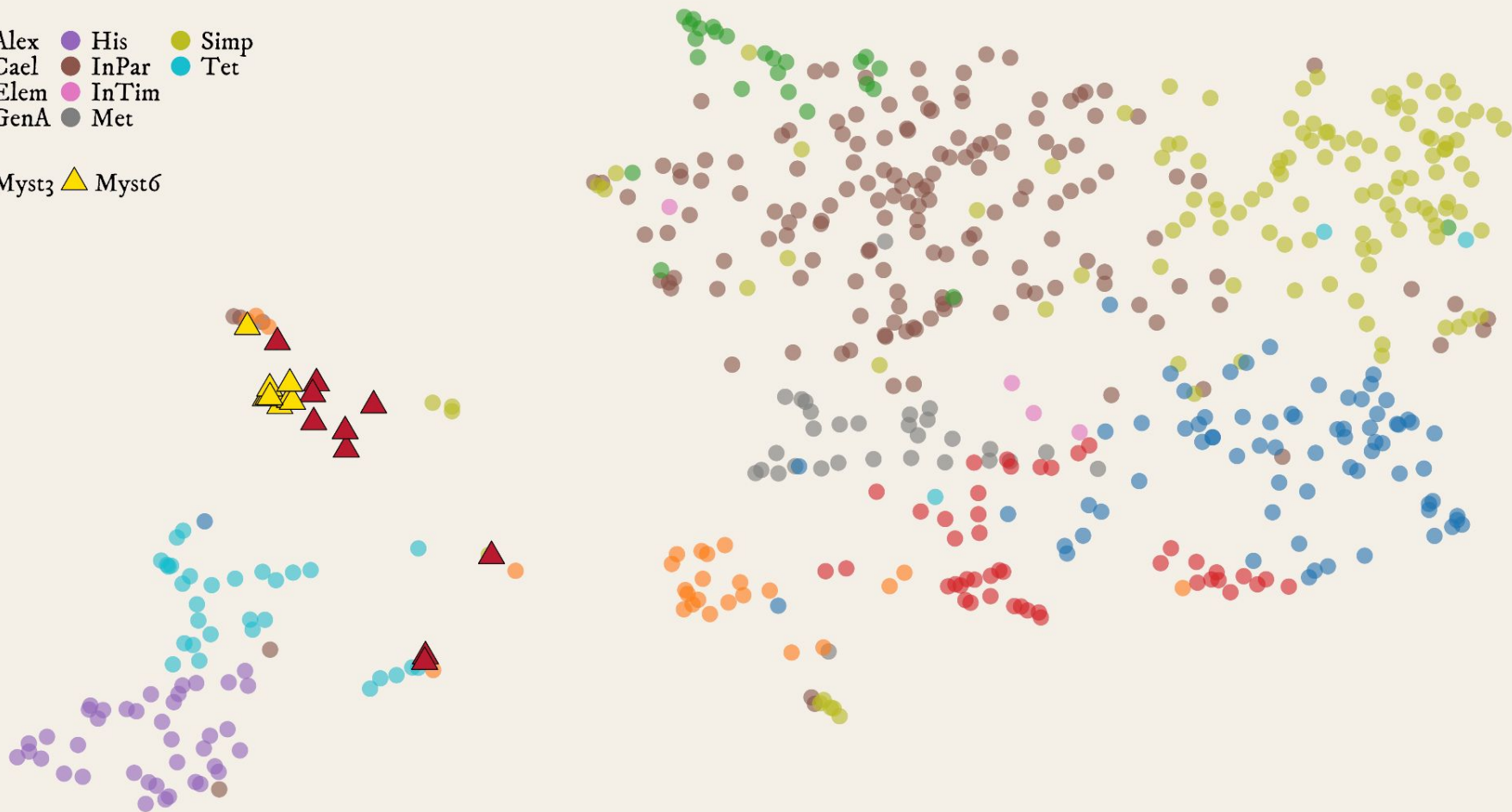
All Works (plus Mystery 3 & 6)



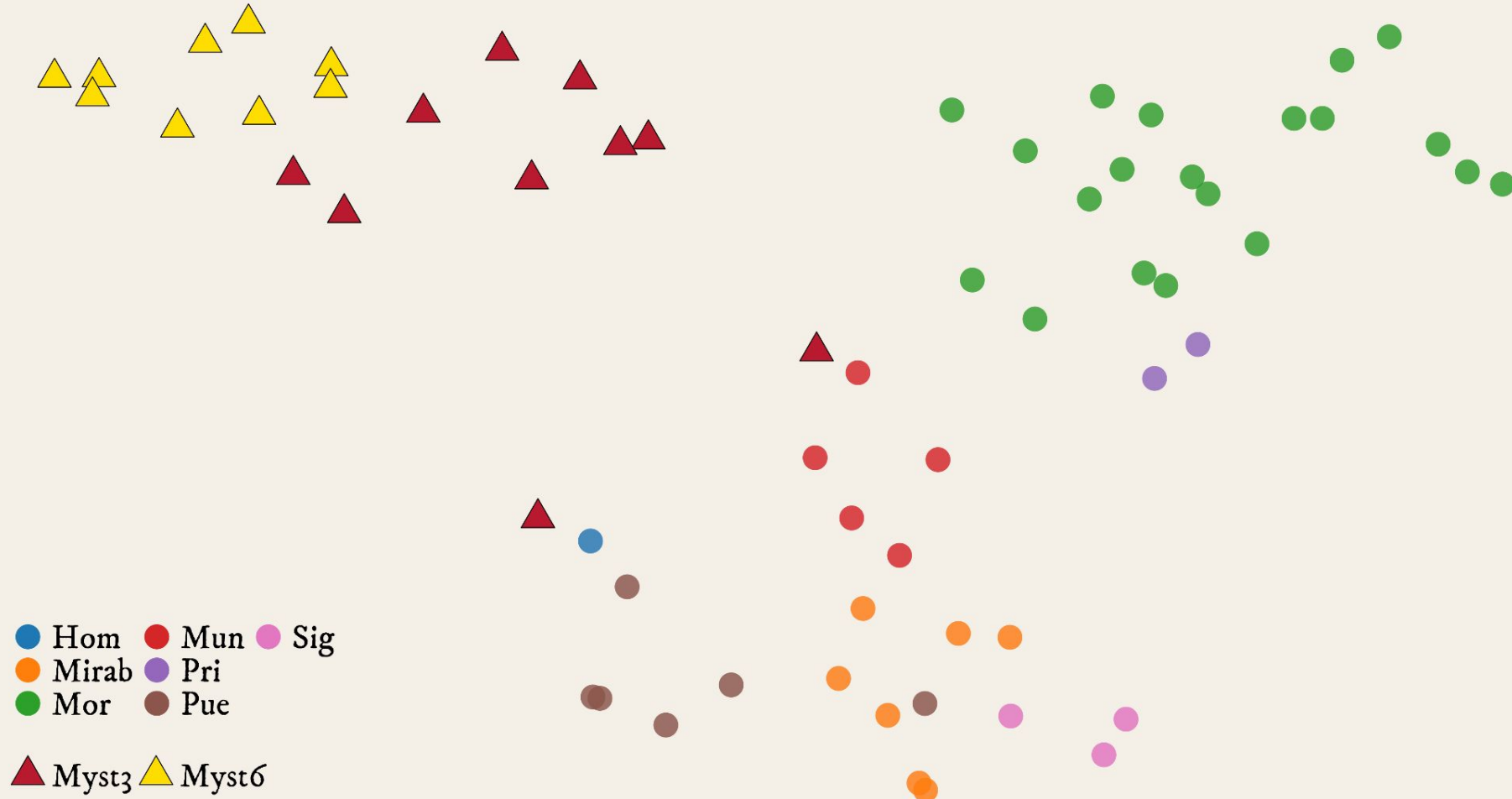
William of Moerbeke

Alex His Simp
Cael InPar Tet
Elem InTim
GenA Met

Myst3 Myst6



Bartholomew of Messina

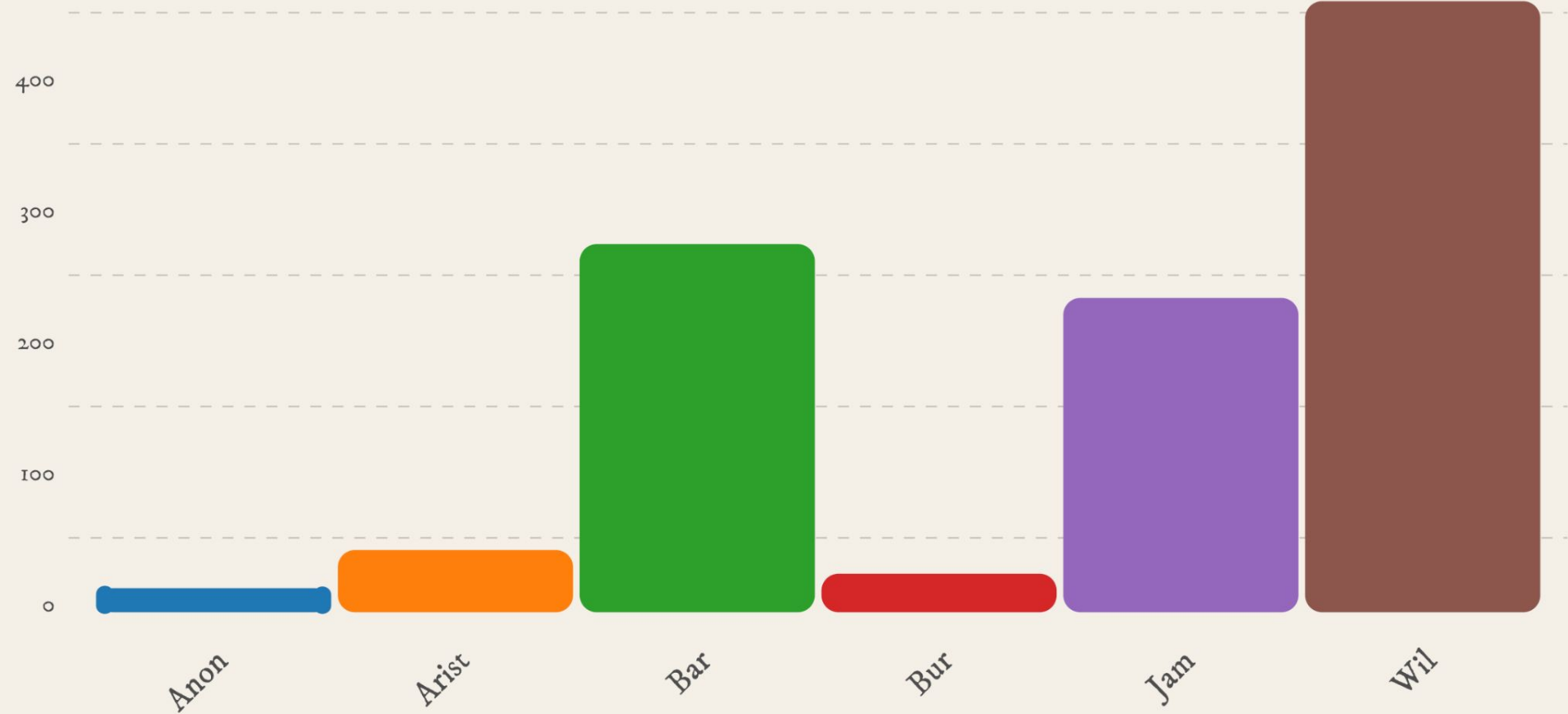


Bootstrap Decision Trees

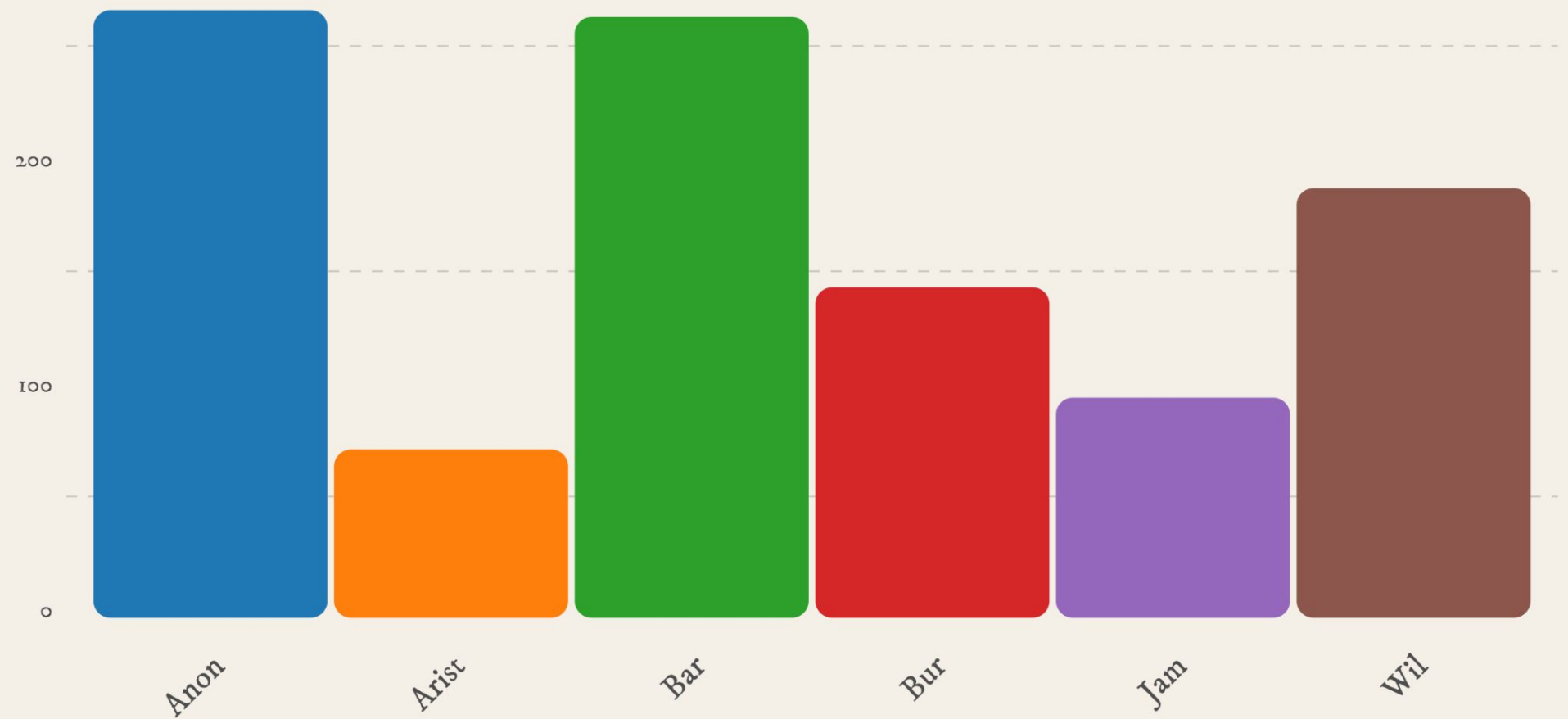
Methods

- 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)



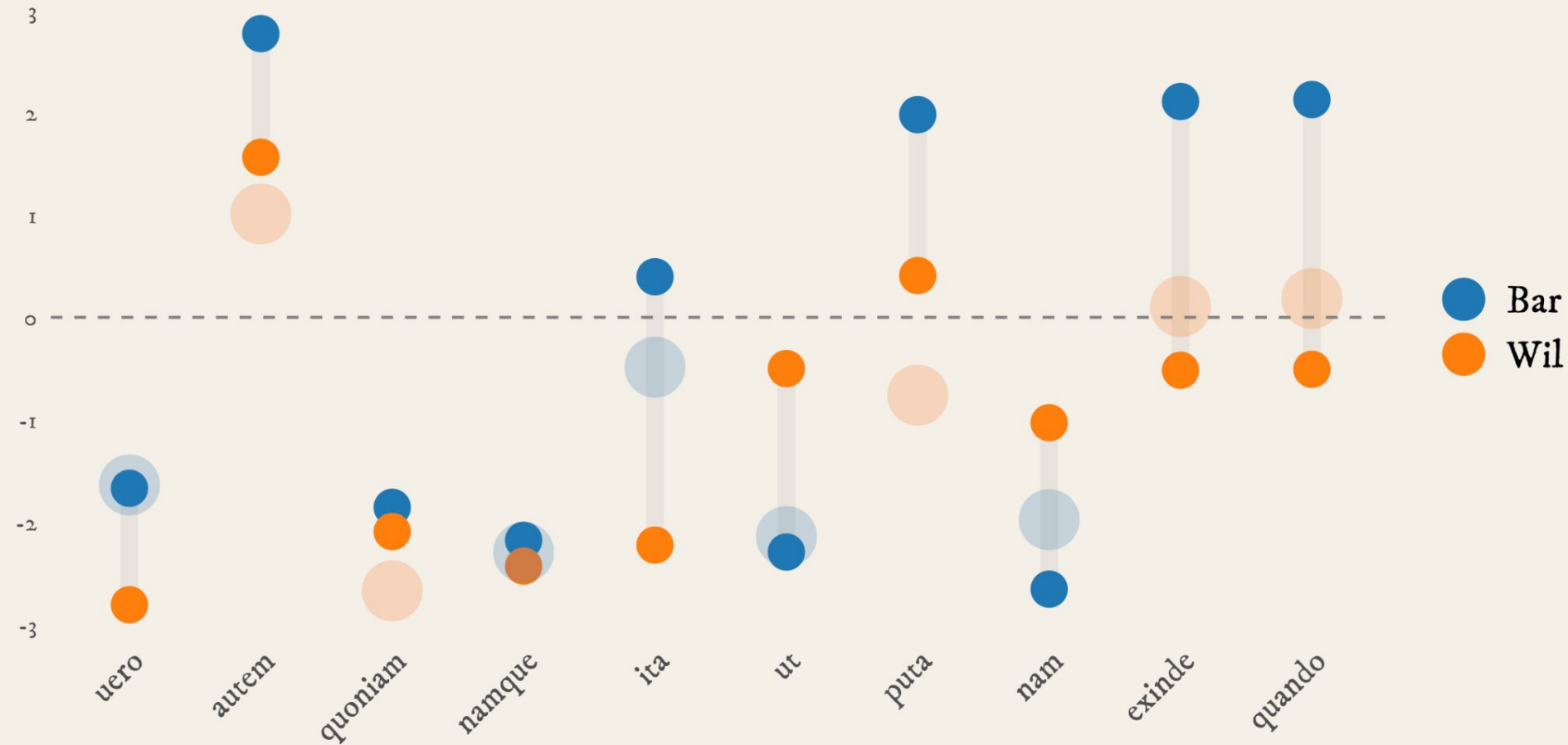
Key

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!