



THE “TONGUEPRINT” AS LANGUAGE IDENTIFICATION TOOL: ELABORATING THE PROOF-OF-CONCEPT

Raf van Rooy, Flavio Massimiliano Cecchini, Isabelle Maes

Presented at UniDive 2nd general meeting, University of Naples “L’Orientale”, 8–9 February 2024 (<https://unidive.lisn.upsaclay.fr/>)



KU LEUVEN

The ERASMOS project

From Chrysoloras’ Latin to Erasmus’ Greek: Renaissance classical bilingualism as a European phenomenon (1397 – 1536)



Έρασμος Ροτερόδαμος (1466 – 1536)

When was the great age of Latin-Greek bilingualism?

The ERASMOS project (ERC grant) argues that it was not in antiquity, but rather during the **Renaissance**, in particular the **long 15th century**:

1397 – the Byzantine scholar Chrysoloras starts teaching Greek in Florence

1536 – Erasmus of Rotterdam dies

In Europe, Latin-Greek bilingualism shaped the cultural identity of the **Republic of Letters**, interacting also with Arabic and Hebrew, and vernacular languages of the time, such as Italian and Dutch.

It was expressed, among other things, through

- Latin-Greek translations
- bilingual compositions, including letters, poems, and more
- code-switching both in texts and oral discourse

The corpus of bilingual texts is vast!

⇒ The **ERASMOS** project aims to organise it and systematically **analyse** the deployment of the two languages in Renaissance literature!

For reference: <https://research.kuleuven.be/EU/p/he/p1/erc/erasmos>

Relevant working groups: WG1, WG3, WG4

The tongueprint

It is the main goal of the project: a **pipeline** to assess the bilingual (or even trilingual, quadrilingual...) make-up of available texts.

The following **tasks** will be part of it:

- detection of **alphabets/fonts**
- optical character recognition (**OCR**) of the text
- detection of the **language(s)** used
- calculation of number, proportion and distribution of languages in terms of **words**
- calculation of the **density** of code-switches and their average word lengths
- analysis of the **level** of code-switches (within or between sentences, paragraphs,...)
- **visualisation** of the data

⇒ The tongueprint will be applied both to **ancient texts** (e.g. Cicero, Suetonius, Martial...) and to **Renaissance texts** (from Chrysoloras to Erasmus), so as to compare the different expressions of bilingualism.

⇒ Beyond Greek and Latin, the tongueprint will be sensitive also to other languages, especially **Hebrew** as a learned one, beside the **vernacular** languages of the time.

⇒ A part of the more refined analyses will involve **morphosyntactic** investigations: the preferred choice to this end will be the formalism of the **Universal Dependencies** project, which already includes treebanks of Latin, Greek, and with code-switching.

Objectives

The main planned **outputs** of the ERASMOS project will be:

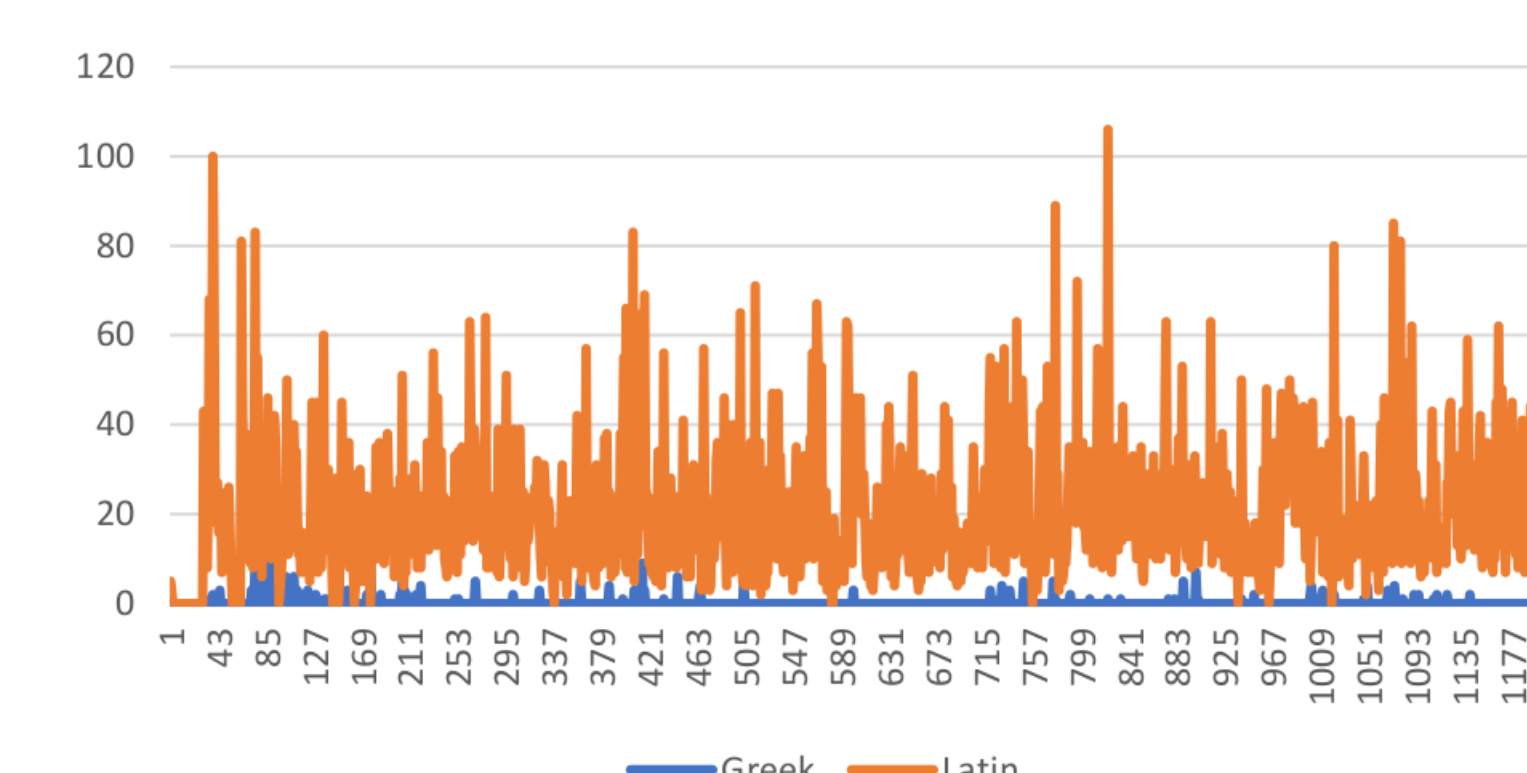
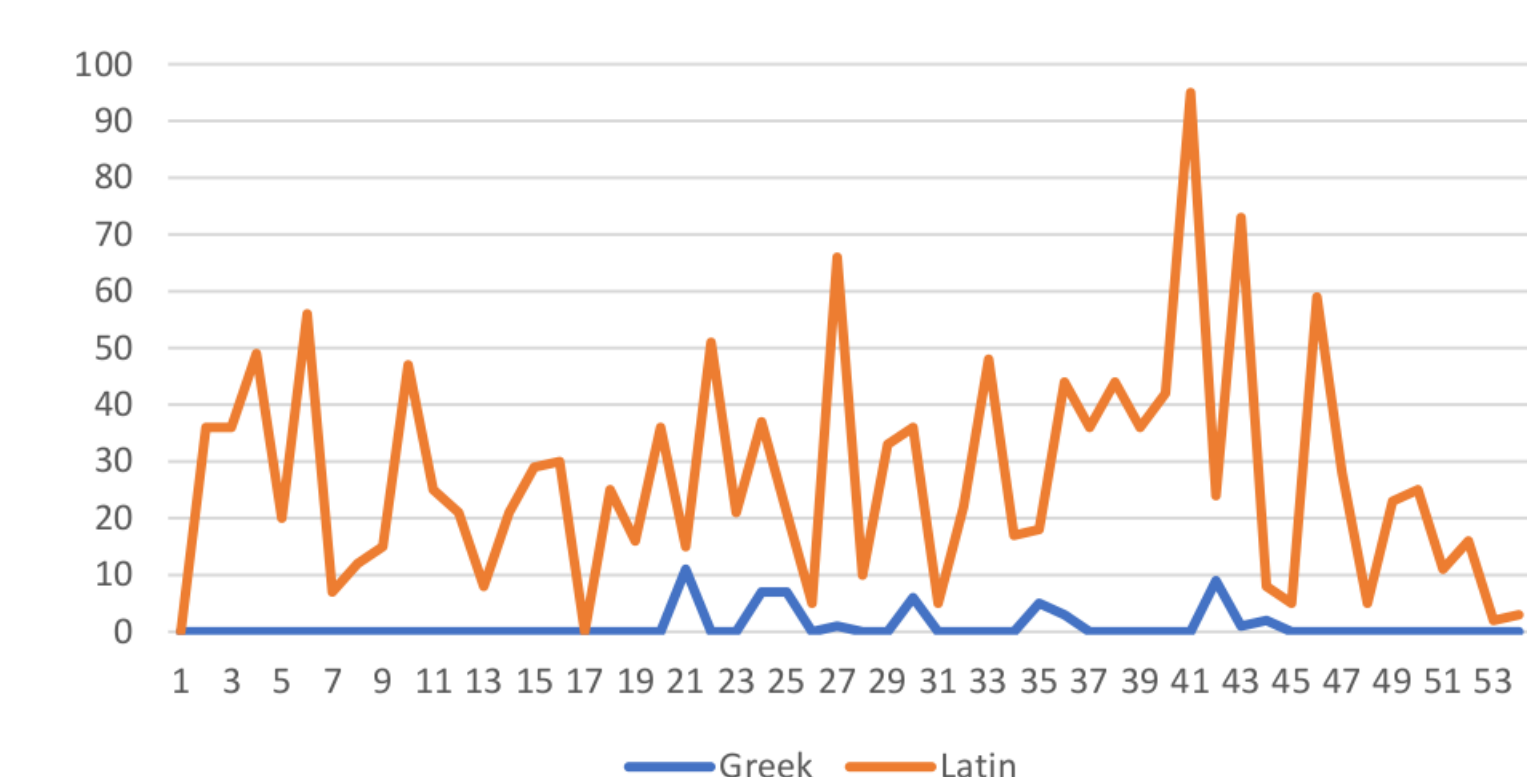
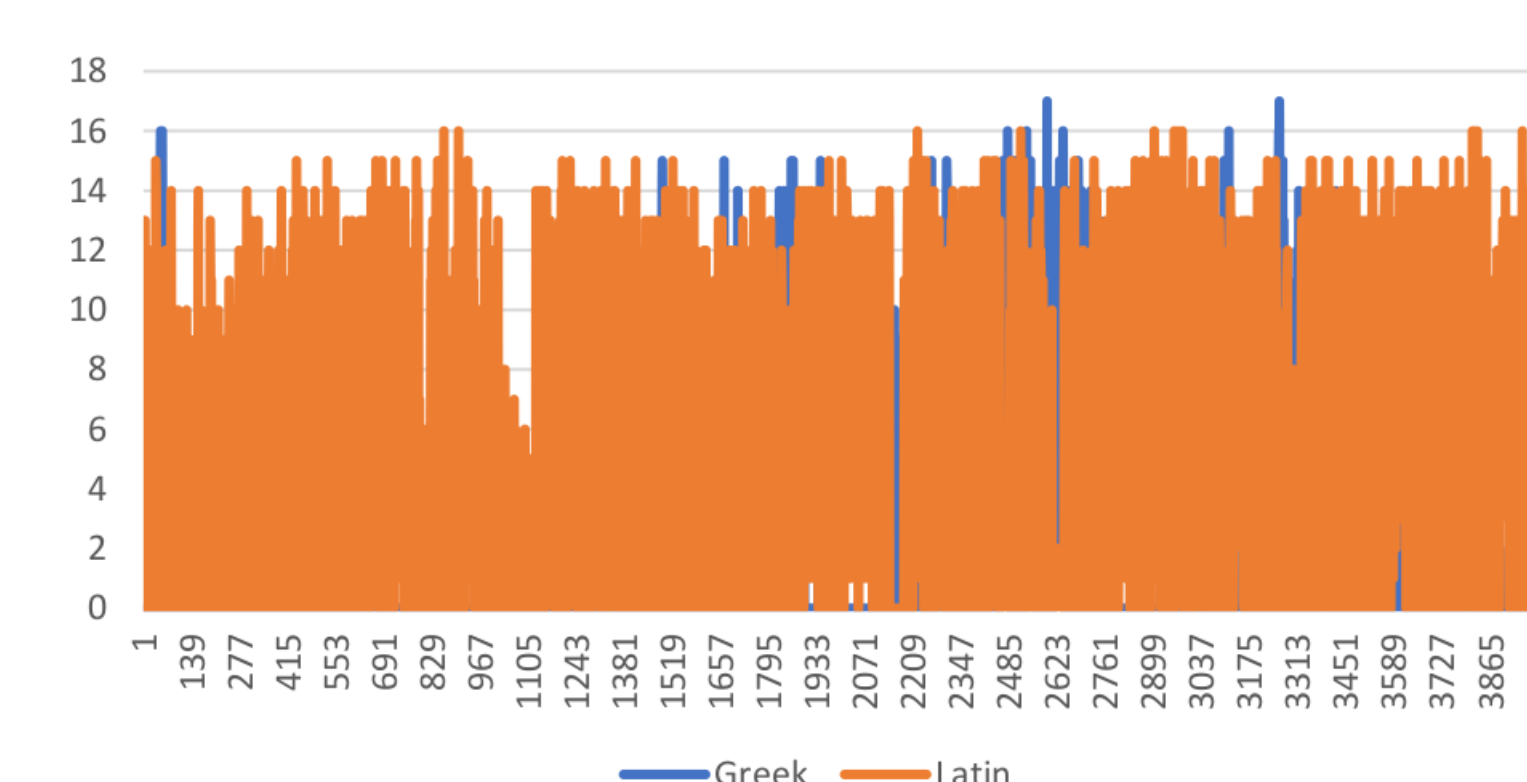
- a **tongueprint pipeline** to analyse and describe the Latin-Greek bilingualism through the ages
- an open-access **corpus** of Erasmus’ oeuvre, with tongueprint analyses for all his works

More broadly, the challenge of the ERASMOS project will be to deal with, highlight and **prioritise multilingualism** in a field where it is mostly ignored, especially for ancient languages of Europe, usually each taken on its own. All **suggestions** are welcome from the community at the start of this endeavour!

Code-switching: preliminary statistics

Comparison of the distribution of **Latin vs. Greek** in three Renaissance writings:

1. the autobiographic journal of Girolamo Aleandro (1480 – 1542)
2. Aleandro’s letter to Erasmus (1512)
3. Erasmus’ *Praise of Folly* (1511/1532)



(from Raf Van Rooy and Wouter Merceis. 2022. *The art of code-switching: Toward a “tongueprint” of multilingual literary personas in Erasmus’ Praise of Folly and Aleandro’s journal?* Leuven Working Papers in Linguistics, 9:1–16)

