THE QUEST FOR LANGUAGE UNIVERSALS: DIACHRONIC TRENDS IN LATIN AND ANCIENT GREEK

Paola Merlo and Kristina Gulordava
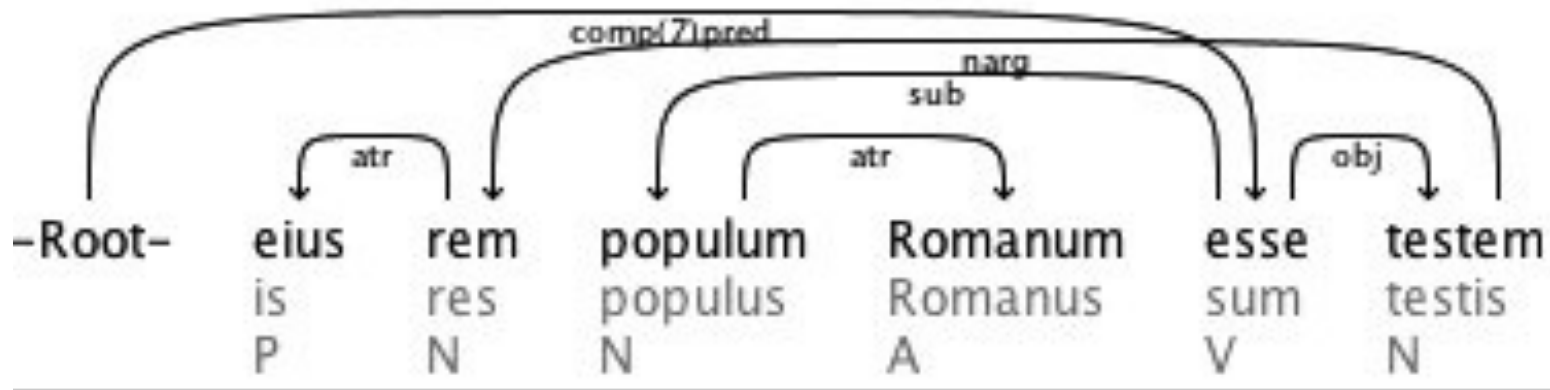
Department of Linguistics

Digital Humanities Short Talks

LANGUAGE UNIVERSALS

- One of the main goals of modern linguistics is to document and study those properties common to all languages, despite languages apparent diversity.

- Current computational methods rely on large repositories of annotated data for many languages, to verify claims of universality on a large scale.
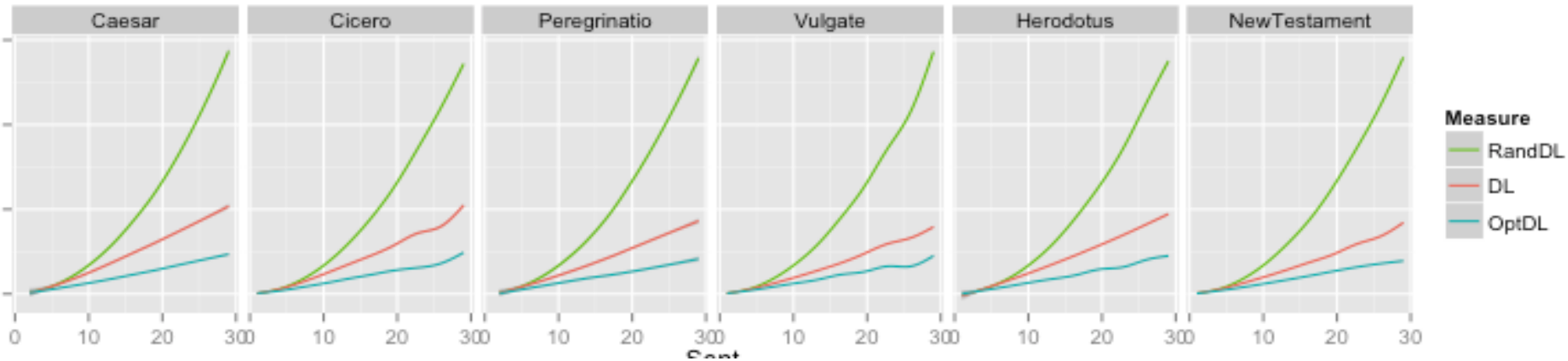
DEPENDENCY TREEBANKS

- We have at the moment annotated data for many languages of many language families. We also have processing tools, such as part-of-speech taggers and syntactic analysers.

WORD ORDER UNIVERSALS

- Word order: some languages show Subject, Object, Verb order in the sentence (German), others show Subject, Verb, Object (French). Some languages have order Adjective Noun, others have Noun Adjective.

- Why do orders differ? Why are some orders more frequent than others?

- Latin and Ancient Greek have large annotated data set at different time periods and the right linguistic properties.

WORD ORDER AND LENGTH OF DEPENDENCIES

Word orders are subject to dependency length minimisation.

Over time, languages tend towards those orders that minimise dependencies.

COMPUTATIONAL LINGUISTICS AND DIGITAL HUMANITIES

- **Need**: The scale of the work requires computational methods and computational resources.

- **Novelty**: We are asking quantitative questions that have been introduced by the existence of large-scale resources.

- **Added-value**:
  We have a formal definition of what counts as 'free word order' and its diachronic trends.

  We develop processing tools for dozens of languages.