A Treebank-based Approach to the *Supprema Constructio* in Dante's Latin Works

Flavio Massimiliano Cecchini, **Giulia Pedonese**
{flavio.cecchini,giulia.pedonese}@unicatt.it

LT4HALA 2022 | Marseille, France | June 25th 2022

*§6 Hunc gradum **constructionis** excellentissimum nominamus, et hic est quem querimus cum **suprema** venemur, ut dictum est. Hoc solum illustres cantiones inveniuntur contexte, ut Gerardus … Folquetus de Marsilia … Arnaldus Danielis […]*

*§7 Nec mireris, lector, de tot reductis autoribus ad memoriam: non enim hanc quam **suppremam** vocamus **constructionem** nisi per huiusmodi exempla possumus indicare. Et fortassis utilissimum foret ad illam habituandam regulatos vidisse poetas, **Virgilium** videlicet, **Ovidium Metamorfoseos**, **Statium** atque **Lucanum**, nec non alios qui nisi sunt altissimas prosas, ut **Titum Livium**, **Plinium**, **Frontinum**, **Paulum Orosium** et multos alios quos amica sollicitudo nos visitare invitat.*

---

*§6 This is the **degree of construction** that I call **most excellent**, and this is what we are looking for when we hunt the best, as I said. Illustrious canzoni are composed using this type of construction alone, as in this one by Giraut … Folquet de Marselha … Arnaut Daniel […]*

*§7 Nor should you be surprised, reader, if so many authorities are recalled to your memory here; for I could not make clear what I mean by the **supreme degree of construction** other than by providing examples of this kind. And perhaps it would be most useful, in order to make the practice of such constructions habitual, to read the poets who respect the rules, namely **Virgil**, the **Ovid of the Metamorphoses**, **Statius**, and **Lucan**, as well as others who have written excellent prose, such as **Livy**, **Pliny**, **Frontinus**, **Paulus Orosius**, and many others whom an affectionate interest invites us to consult.*

What is this **supreme degree of construction**?

▶ the perfect **syntactic structure** to write poetry in **Italian** in the highest style possible

▶ appliable to both **Romance** languages (Old French, Old Provençal and Italian) and **Latin**

▶ inspired by both Latin **poetry and prose**

Dante participated in the XIV's century **diglossia**

Italian works:

► **Rime** Rhymes
► **Vita Nova** The New Life
► **Convivio** The Banquet
► **Commedia** Divine Comedy

Latin works:

► **Epistulae** Letters
► **De vulgari eloquentia** On vernacular eloquence
► **De Monarchia** On Monarchy
► **Eglogae** Eclogues
► **Questio de aqua et terra** An inquiry about water and earth

Style or **syntactic** structures, or both?

gradus insipidus flavourless degree

$\longrightarrow$ *Petrus amat multum dominam Bertam*
'Peter loves Miss Bertha a lot'
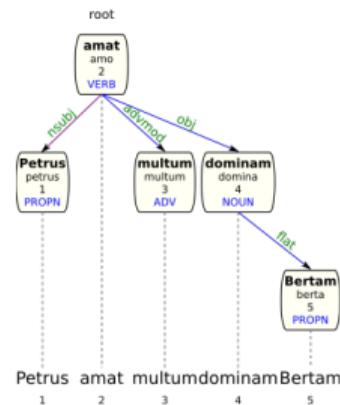
Style or **syntactic** structures, or both?

gradus sapidus flavoured degree

$\longrightarrow$ *Piget me cunctis pietate maiorem, quicunque in exilio tabescentes patriam tantum sompniando revisunt*

'I am stricken with sorrow more than most, for whomever drags out his life in exile, revisiting his native land only in dreams'
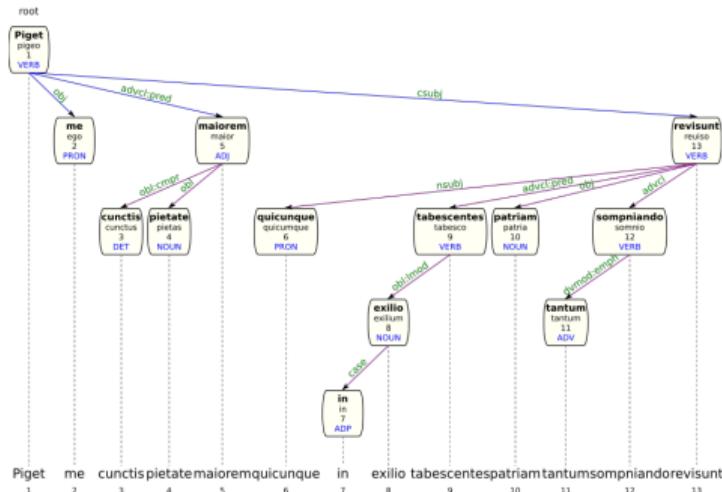
Style or **syntactic** structures, or both?

gradus sapidus et venustus flavoured and graceful degree

⟶ *Laudabilis discretio marchionis Estensis, et sua magnificentia preparata, cunctis illum facit esse dilectum*

'The laudable discretion of the Marquis of Este, and his widely displayed generosity, make him beloved of all'
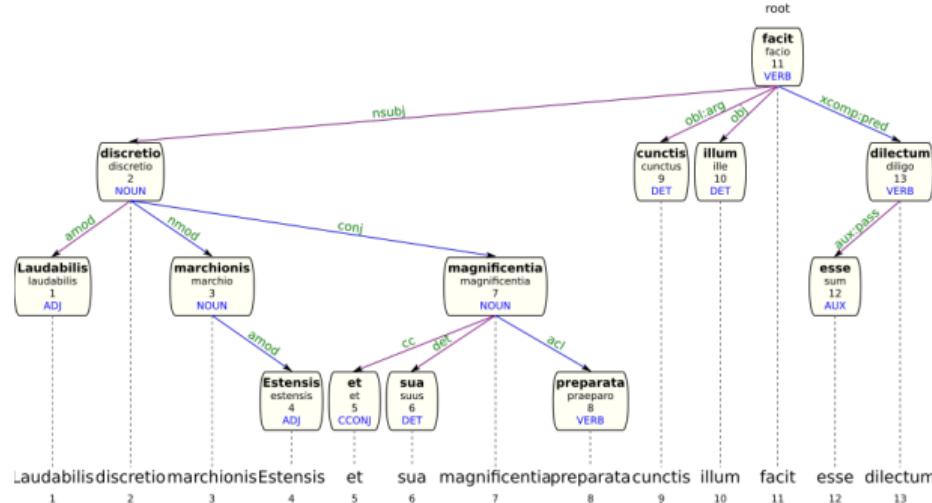
Style or **syntactic** structures, or both?

**gradus sapidus et venustus etiam et excelsus** flavoured and graceful and also striking degree

⟶ *Eiecta maxima parte florum de sinu tuo, Florentia, nequicquam Trinacriam Totila secundus adivit*

'The greater part of your flowers, o Florence, having been snatched from your breast, the second Totila advanced in vain towards Trinacria'
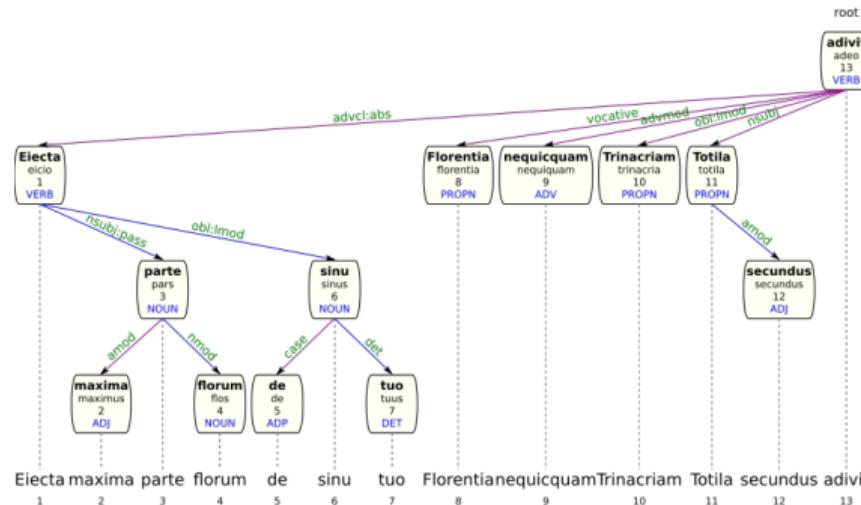
*§8 Subsistant igitur **ignorantie** sectatores **Guittonem Aretinum** et quosdam alios extollentes, nunquam in vocabulis atque **constructione** plebescere desuetos!*

*§8 So let the devotees of **ignorance** cease to cry up **Guittone d'Arezzo** and others like him, for never, in either vocabulary or **construction**, have they been anything but commonplace.*
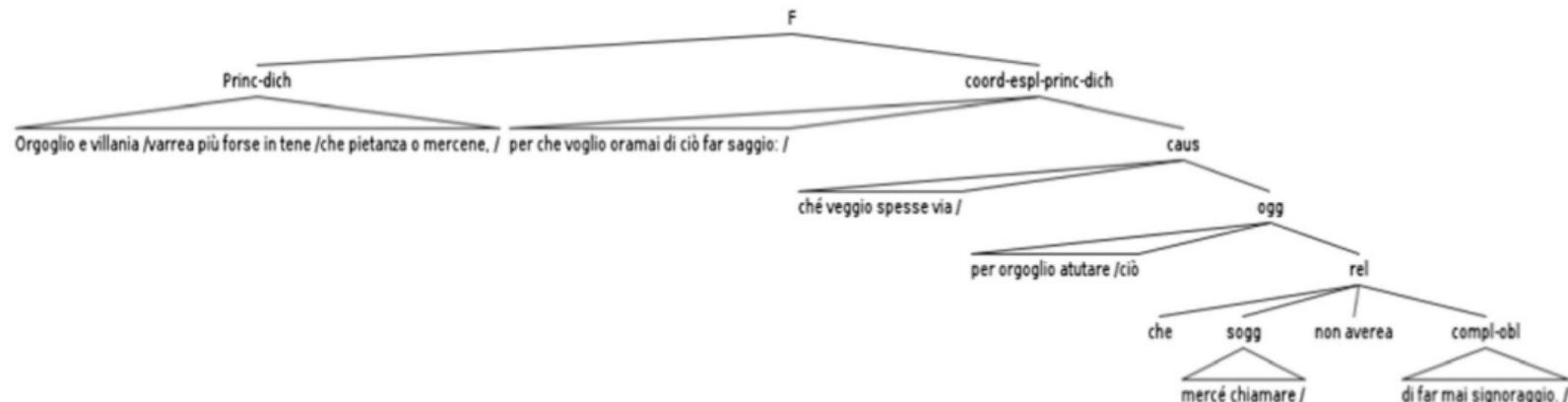
What do we know about Dante's syntax?
The "state of the syntax"

IRCSE 9

Most linguistic studies and language resources have been focusing
on Dante's **vernacular (Italian)** works

► **DanteSearch**

⟶ all Latin & vernacular works by Dante, lemmatised and annotated for morphology and partially (Rhymes, Banquet, Comedy) for syntax with an own tagset

► **Vocabolario Dantesco Latino** Latin Vocabulary of Dante's works

⟶ with the goal of providing the first systematic study of Dante's Latin lexicon through Classical and Medieval Latin corpora

No description of Dante's Latin syntax has been attempted despite the **importance** of such analysis in the relationship between the author and his sources.

One of five corpora of Latin in **Universal Dependencies**

▶ All five (known) **Latin works** by Dante: *Epistulae*, *De vulgari eloquentia*, *De Monarchia*, *Eglogae*, *Questio de aqua et terra*

▶ 1 721 sentences for **55 503 tokens**

▶ **All levels** of annotation (lemmas, parts of speech, morphology, syntax)

▶ Syntax **manually annotated** from scratch by a group of domain experts

▶ based on the **DanteSearch** corpus (University of Pisa)

Cecchini, Flavio Massimiliano, Rachele Sprugnoli, Giovanni Moretti, and Marco Passarotti. "UDante: First Steps Towards the Universal Dependencies Tree- bank of Dante's Latin Works." In Proceedings of the Seventh Italian Con- ference on Computational Linguistics (CLiC-it 2020, Bologna, Italy, March 1–3 2021)

In the **UD framework**, data with regard to authors cited by Dante is **lacking**

PROIEL and **Perseus** contain Classical texts, but…
- … of those cited by Dante, only **Vergil** and **Ovid** are found (in Perseus)
- … only **book I of Metamorphoses** and **book VI of Aeneid** are present
- … of which many sentences are **missing**
- … and **Perseus** has **"neglected"** status as of UD v2.10
    - ! `http://quest.ms.mff.cuni.cz/udvalidator/cgi-bin/unidep/validation-report.pl`

Automated annotation for these texts is still **not good enough**:

$\Rightarrow$ **U**nlabeled **A**ttachment **S**core for UDPipe on a sample of these texts: **40.2%**

For now, we have to **limit** the study to a comparison between **Dante's works**

> Tavoni, Mirko, and Emmanuele Chersoni.
> "Ipotesi d'interpretazione della «supprema constructio» (*De vulgari eloquentia* II vi)."
> *Studi di grammatica italiana*, XXXI–XXXII (2013): 131–158

▶ …**attempts** to formally describe the *supprema constructio* by extending the syntactic annotation of Dante's Italian works in **DanteSearch**

▶ …**calculates** an **asymmetry** index and a **(dis)homogeneity** index and applies it to the French, Provençal and Italian poems quoted by Dante

▶ …**links** Dante's definition of *supprema constructio* to the idea of **syntactic balance**: pattern characterized by **low depth and symmetry**

▶ …**points** out that the analysis should be extended to the other cited Latin authors

⟶ Vergil, Ovid's Metamorphoses, Statius, Lucan, Livy, Pliny, Frontinus and Paulus Orosius

We try to propose a new, more **synthetic and "universal"** measure of syntactic balance

⇒ The barycentre of a graph is the set of nodes whose **sum of distances** from each other node (total distance) is **minimal**
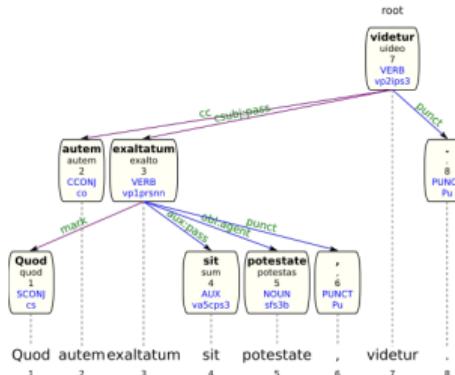

De vulgari eloquentia I *xvii* 4

**Root** : linguistic ⇔ **Barycentre** : topological
The barycentre represents the **"main branching point"** with respect to the sentence **"hub"**

Results on Dante's works
Two antipodes
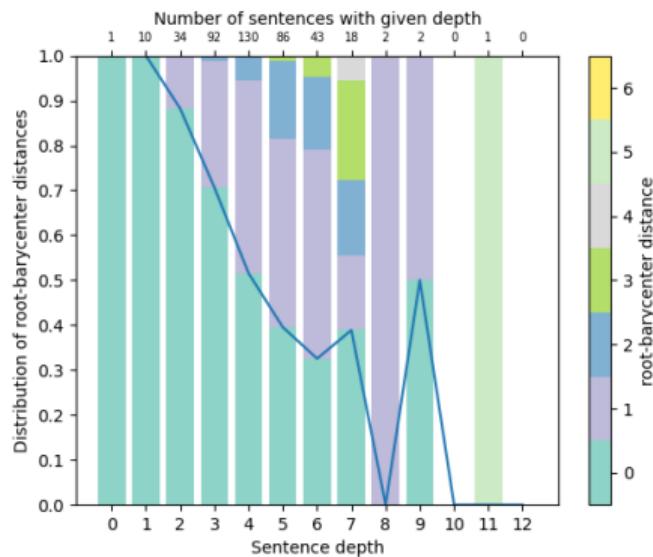
De vulgari eloquentia
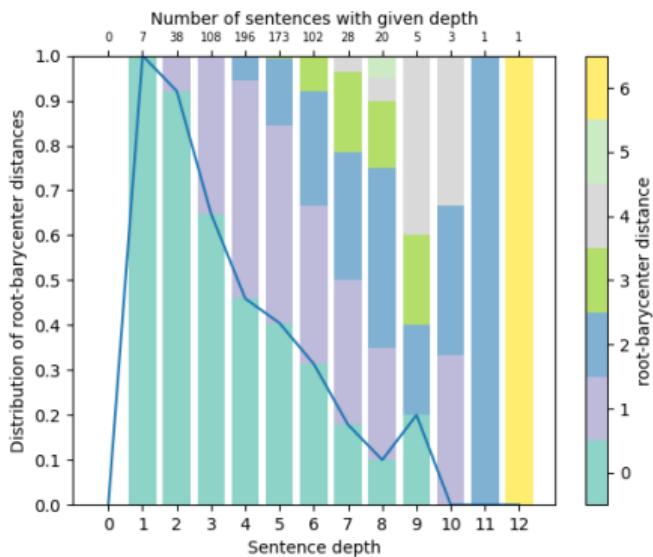
De Monarchia
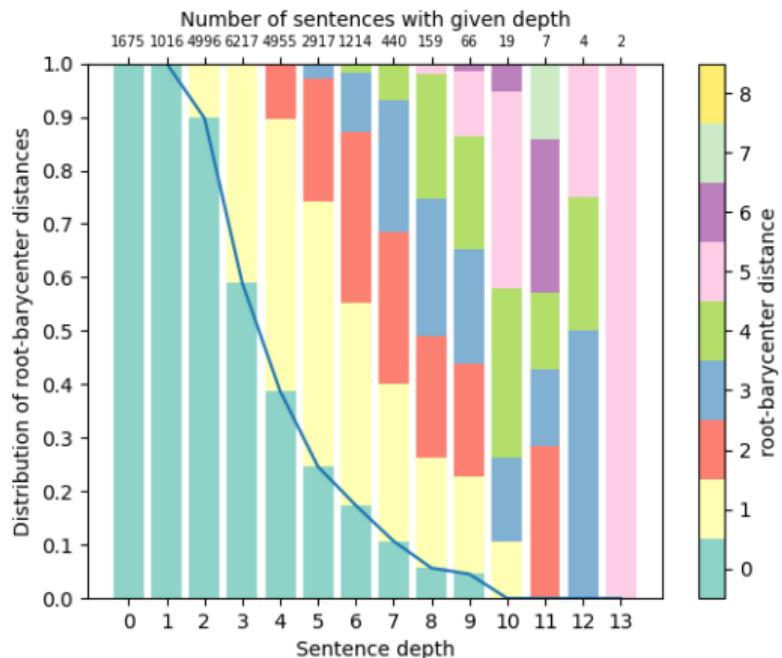
We observe some **tendencies**, but cannot yet come to definitive conclusion:
the nature of *supprema constructio* stays **elusive**

▶ Obtain or create more Classical Latin **syntactic data**

▶ Study the behaviour of root-barycentre delta for **other languages**

▶ Expand on other factors like **non projectivity**

A **supreme** thank you for your attention!

Flavio Massimiliano **Cecchini** & Giulia **Pedonese**

Università Cattolica del Sacro Cuore di Milano
CIRCSE research centre
Largo Gemelli 1, 20123 Milano

✉ {flavio.cecchini,giulia.pedonese}@unicatt.it

○ https://github.com/Stormur/DanteSuppremaConstructio
○ https://github.com/CIRCSE
🌐 https://centridiricerca.unicatt.it/circse-home
🌐 https://itreebank.marginalia.it/
🌐 https://lila-erc.eu