


Sunoikisis Digital Classics, Fall 2020  
November 19, 2020



# Linguistically annotated corpus of documentary papyri: PapyGreek (Paratypa, Sematia)

Marja Vierros (University of Helsinki)



European Research Council  
Established by the European Commission

Digital Grammar of Greek Documentary  
Papyri (PapyGreek)  
ERC Starting Grant, No. 758481

---

# Results from the Genitive Absolute queries

Marja Vierros and Polina Yordanova. (Forthcoming). Querying syntactic constructions in Ancient Greek parsed corpora: a case study on the genitive absolute in literature and documentary papyri. In Novokhatko, Anna, Stylianos Chronopoulos and Felix Maier (eds) *Classics@ CHS*



## GA in literary corpora (1)

<i>Genre/ writer</i>	<i>Corpus: Writers</i>	<i>sentences with GA</i>	<i>GA / Number of sentences</i>
Epic and drama	AGDT: Hesiod, Aeschylus, Sophocles, Ps-Homer	1,5%	1,6%
Plato	Gorman: Apology AGDT: Euthyphro	2,4%	2,7%
Xenophon	Gorman	6,2%	7,3%
Aristotle	Gorman	4%	4,8%
Rhetorics	Gorman: Antiphon, Aeschines, Lysias 5th/4th BCE	6,5%	6,7%
	Gorman: Demosthenes 4th BCE	10,5%	13,2%



## GA in literary corpora (2)

Early historical prose	Gorman + AGDT: Herodotus	9%	10%
	Gorman + AGDT: Thucydides	11%	13%
Later historical prose	Gorman + AGDT: Polybius (2nd BC)	18%	22%
	Gorman + AGDT: Diodorus Siculus (1st BCE)	25%	31%
	Gorman: Dionysius Halicarnassus (1st BCE)	13%	15%
	Gorman: Josephus (1st CE)	11%	13%
	Gorman + AGDT: Plutarch ((1st/2nd CE)	17%	21%
	Gorman: Appian (2nd CE)	21%	26%

# GA in documentary papyri (PapyGreek Sematia)



<i>Text type</i>	<i>Text count</i>	<i>Texts with GA (% of texts within type)</i>	<i>GAs within type (% of GAs within corpus)</i>
Letter	153	17 (11%)	24 (19%)
Petition	44	37 (84%)	91 (71%)
Contract	21	5 (24%)	5 (4%)
Other types	7	4 (57%)	7 (5%)
Text type not defined	25	1 (4%)	1 (0.8%)



## Summary on the GA (1)

- it was one element in the stylistic repertoire of especially historical narrative writing, more popular in the Hellenistic and Roman periods
- in papyri most common in petitions, which mainly consist of a narrative of some event when the petitioner has been maltreated
- everyday letter writing had no place for the GA (more official letters could use it)



## Summary on the GA (2)

- The suggestions that in Koine Greek the GA's are more often **not “absolute”** (i.e. the subject of GA plays syntactic role in the matrix clause; extreme case: subject of the matrix clause and subject of GA coincide, see below):
  - cannot fully be searched computationally
  - qualitative analysis from small sample of papyri reveal that ca. 20% were not absolute

ἀπελθόντος μου      ἀπὸ σοῦ κατέλαβον τοὺς γεωργο[ὺς]... ἀνακεχωρηκότας... (P. Cair. Zen. 2 59245, 1)

apelthontos mou      apo sou **katelabon** tous georgous... anakechorekotas

PTC:gen.m.      PR.1sg.gen.m      discover:1sg

*After I left you, I discovered that the peasants ... had fled...*

---

# PapyGreek project and platform







## DIGITAL GRAMMAR OF GREEK DOCUMENTARY PAPYRI

ERC starting Grant funded project 2018–2023

Marja Vierros, PI  
(morphosyntax)

Sonja Dahlgren  
postdoc  
(phonology)

Erik Henriksson  
(developing  
infrastructure)

Polina Yordanova  
PhD candidate  
(word order, treebanking)

Research assistants (shorter time periods)

Arttu Alaranta, Petri Lahtinen, Jamie Vesterinen, Iida Huitula, Sari Kock

Project website: <https://www.helsinki.fi/en/researchgroups/digital-grammar-of-greek-documentary-papyri>

Platform: <https://papygreek.hum.helsinki.fi/>





# PapyGreek platform

(<https://papygreek.hum.helsinki.fi/>)

- Variation: Paratypa, search tool for orthographic variation + treebanks
- Sematia: Layering and preprocessing the EpiDoc XML texts for treebanking
- Integrated Arethusa annotation tool (originally developed for the Perseids Project (Tufts)
  - [Guidelines 2.0](#) (G. Celano) < [Guidelines 1.1](#). (D. Bamman & G. Crane 2008)
  - [PapyGreek guidelines for papyrological structures and for keeping the corpus uniform](#)
- Addition of metadata for treebanked texts (author / writer / addressee, text type)
- In progress!

# Paratypa

- Dynamic search tool for orthographic variations in the papyri
- Queries differences between the original text and editorial markup (ORIG vs. REGularized)
- The whole Duke Databank of Documentary Papyri -corpus, updated regularly
- Preceding and following context can be defined (immediate or further; word/line/sentence initial or final etc.)
- E.g. <λιπ> written instead of <λειπ> = **addition** of <ε> by the editor

## Variations

Replacements Additions Deletions Unchanged

Greek Latin

☐ Diacritic sensitive

← BEFORE →

← AFTER →

← λ →

→ ε ←

→ ιπ ←

Showing 1 to 50 of 87 entries

Filter:

Previous 1 2 Next

Original ↕	Regularized ↕	Document ↕	≤ ↕	≥ ↕	🌐 ↕	Textpart ↕	Line ↕
_ἀδιαλίπτως	_ἀδιαλείπτως	<a href="#">bgu.8.1854.xml</a>	-45	-44	Herakleopolites		4
ἀδιαλιπ	ἀδιαλείπτως	<a href="#">sb.14.12168.xml</a>	130	130	Oxyrhynchites		19-20
ἀδιάλιπτόν	ἀδιάλείπτόν	<a href="#">p.lond.5.1658.xml</a>	301	400	unbekannt	r	6
ἀδιαλίπτως	ἀδιαλείπτως	<a href="#">p.panop.7.xml</a>	338	339	Panopolis		5
ἀδιαλίπτως	ἀδιαλείπτως	<a href="#">psi.9.1037.xml</a>	301	301	Oxyrhynchos	r	17
ἀδιαλίπτως	ἀδιαλείπτως	<a href="#">sb.14.12088.xml</a>	346	346	Oxyrhynchos		13
ἀδιαλίπτως	ἀδιαλείπτως	<a href="#">p.oxy.1.83dupl.xml</a>	328	328	Oxyrhynchos		12-13
ἀδιαλίπτως	ἀδιαλείπτως	<a href="#">p.oxy.47.3354.xml</a>	257	257	Nomu (Oxyrhynchites)	r	19

## Comparing to the correct spellings

- <λιπ> written correctly (not corrected by editor)

Replacements Additions Deletions Unchanged							
Greek Latin							
<input type="checkbox"/> Diacritic sensitive							
< λιπ >							
Showing 51 to 100 of 1,205 entries Filter:							
Previous 1 2 3 4 5 ... 25 Next							
Original	Regularized	Document	≤	≥		Textpart	Line
ἀπολίπω		p.oxy.3.494.xml	165	165	Oxyrhynchos		9
ἀπολίπω		p.oxy.66.4533.xml	76	125	Oxyrhynchos		11
ἀπολίπω		p.oxy.3.492.xml	130	130	Oxyrhynchos		13
ἀπολίπω		p.oxy.3.490.xml	124	124	Oxyrhynchos		6
ἀπολίπω		p.oxy.66.4533.xml	76	125	Oxyrhynchos		6
ἀπολίπω		p.oxy.3.492.xml	130	130	Oxyrhynchos		7
ἀπολίπω		p.oxy.3.495.xml	181	184	Oxyrhynchos		6

# Layering the original and regularized text for annotation

SEMATIA

o.claud.2.252.xml

TM: 29672 HGV: 29672 Date: 126 – 175 Place: Tiberiane

View

Edit tokens

Import data

Annotate



Edit annotation

Metadata

Versions

- Two layers will be externally annotated for morphology and syntax (in Arethusa)
- Regularizations highlighted
- Automatically parsed morphology is utilized for regularized part (Leuven-Pedalion data), syntax not yet





1  m1 S-1 Πετενεφώτης |gap:1| Σαραπίωνι



2 τῷ τιμιωτάτῳ  χαίρων  
 χαίρειν .

None  None [0]

2  m1 S-2  ἐπιε  
 ἐπεὶ

2-3 λέγουσιν

3  ἔχων σε ἐπιστολὰς  ὑμῶν  
 ἔχειν  ἡμῶν

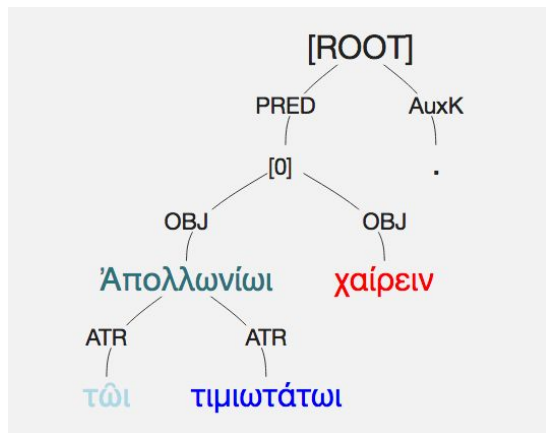
4 ἀπὸ Αἰγύπτου , εὖ  ποιήσεις , ἐὰν  
 ποιήσεις

5 ἔχης μου ἐπιστόλια ,

5-6 πέμψον

## Regularized

[0] |gap:7| Ἀπολλωνίῳ τῷ τιμιωτάτῳ χαίρειν .

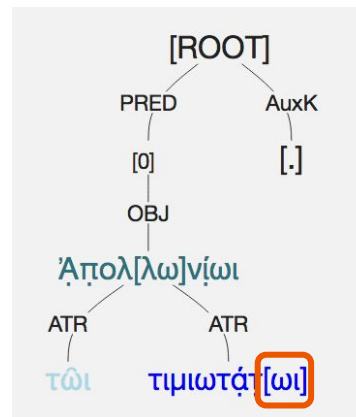


✓ τίμιος  
adj.sg.masc dat.sup

a-s---mds

## Original

|gap:7| [0] Ἀπολ[λ]ωνίῳ τῷ τιμιωτάτ[ωι] χαί[ρ]ει[ν] [.]



✓ τίμιος  
adj.masc.sup

a-\_-m\_s



Lacunae interpreted; here Apollonios can be annotated, χαίρειν left out

Ἀπολ[λ]ωνίῳ 1-2

✓ Ἀπολλώνιος n-s---md-  
noun.sg.masc.dat

The ending is in a lacuna, not marked as dative in the original layer

---

# Demo

Paratypa Variations

Main question: **genitive and dative variation**

Looking more closely at **second declension singular endings -ου vs. -ω:**