

Download

XML-based transcriptions of medieval Greek manuscripts: Balancing the needs of different end- users

from

Notis Toufexis
nt262@cam.ac.uk
University of Cambridge
Grammar of Medieval Greek project

<http://www.toufexis.info>

Overview of presentation

Download

- Context of overall research
- Clarifications
- More general questions
 - Transcription and edition
- A case study *from*
 - Transcription of a 16th c. manuscript

<http://www.toufexis.info>

Context of research

Download

- The study of the multi-layered system of Medieval Greek
 - Archaic / learned vs. vernacular
 - Register & linguistic variation
 - Koine vs. dialect *from*
 - Linguistic variation
 - Foreign vs. native
 - Romance, oriental elements
- Development of literary and non-literary registers towards Modern Greek

<http://www.toufexis.info>

(Grammar of) "Medieval Greek"

Download

1. = Vernacular Greek, closest approximation to the spoken language
 - Reconstruction with the help of linguistic theory
2. Not denying the reality of the texts, accepting register variation
 - Mixing of vernacular with learned elements and vice versa
3. Medieval period covers a larger time span (1100-1700)
 - "Byzantium after Byzantium"

<http://www.toufexis.info>

“Transcription”

Download

- As accurate as possible presentation of the text preserved in a (manuscript) witness with no or minimal interventions by the editor
 - Use of sigla for abbreviations, ligatures, nomina sacra etc.
 - Level of palaeographical detail matter of choice
 - “Mistakes” remain in the text of the transcription
 - “Problem is, one era's nonsense is another's orthography, it would seem.” (Byzans-L)

Transcription and edition (pre-digital)

- Philologists seem quite reluctant to accept the need for publication of manuscript transcriptions for literary texts
- Strong preference for editions
 - Irrespective of editorial method (e.g. stemmatic, copy-text etc.)
- Transcription considered as a prerequisite for and not equivalent to an edition

<http://www.toufexis.info>

“Edition”

Download

- (As accurate as possible) presentation of the text for the modern reader
 - Mediation of the editor
 - Choice of method (stemmatic, copy-text, best-text)
 - Choice of conventions (spelling, text structure etc.)
 - Documentation of decision making

<http://www.toufexis.info>

Editorial method and transcription

Download

- Stemmatic method
 - Collation of variant readings
 - No need for full transcription of manuscript witnesses
- Copy-text etc. *from*
 - Decision lies with the editor, full transcription of manuscripts possible
 - Based on the level of variation, simple collation of variants possible

<http://www.toufexis.info>

Transcription stays with the editor

Download

- By-product of the edition
- Concealed from end-users of edition
- Users of edition interested in manuscript readings rely on the editor's documentation *from*
 - Apparatus criticus, introduction of edition, critical commentary

<http://www.toufexis.info>

Transcription as Edition

Download

- “Diplomatic method” acceptable for non-literary documents
 - Monastic acts
 - Inscriptions
 - Historical documents
 - etc.

from

<http://www.toufexis.info>

Transcription, digital

Download

- As part of a digital edition
 - The one text and the many texts (Robinson 2000)
- As research tool
 - To facilitate further research
- Digital transcription open for end-users

from

<http://www.toufexis.info>

Digital transcriptions have to take into account the needs of users

- Can we always foresee how a digital transcription will be used?
 - Different interests by different end-users
 - Palaeographers vs. Linguists vs. Non-specialists
- Can we make all happy?
- Are there minimal requirements?
- Is this a new research tool?

<http://www.toufexis.info>

Digital transcription as research tool: possible applications for Medieval Greek

- Composite manuscripts
 - Explore the role of the manuscript scribe in transforming different texts in the same codex
 - Of interest for philologists
- Manuscripts with linguistic interest
 - Autographs of authors
 - Language use by authors as clue for the development of linguistic registers of Greek
 - Of interest for historical linguists

<http://www.toufexis.info>

A case study: Θεματοεπιστολαί

Download

Θεματοεπιστολαί.

- Written 1578 by Theodosios Zygomalas in Constantinople for Martinus Crusius in Tübingen
- Part of Codex Tyb. gr. Mb 19



<http://www.toufexis.info>

Intralingual translation of letter templates

Download

- Martinus Crusius tries to learn the spoken Greek language of his time
 - Orders “teaching material” from Zygomalas
- Zygomalas puts a compilation of letter templates together and translates them in the vernacular
 - Learned and vernacular next to each other by the same author

<http://www.toufexis.info>

Why is this interesting?

Download

- Allows us to draw conclusions on the author's attitude towards language
 - Author with solid classical education (by the standards of this period)
 - Author of other intralingual translations of (legal) texts
 - Source and target language side by side
- Byzantine genre in a new linguistic form
 - Makes the development of registers in the vernacular more transparent

from

<http://www.toufexis.info>

Objectives

Download

- Make the (text of the) manuscript accessible to a broader research community with linguistic interests
 - Online publication
- Interdisciplinary perspective
- Provide with an accurate transcription of the manuscript
 - Conventions (?)
 - Digital facsimiles (?)
- Beyond the transcription, towards an edition
 - Alignment, glossary, commentary etc.

<http://www.toufexis.info>

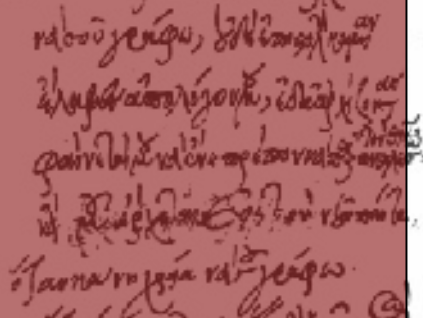
Decisions

Download

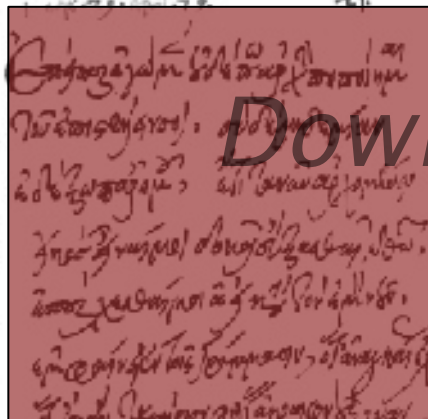
- TEI P4
 - <!ENTITY % TEI.transcr 'INCLUDE' >
 - Encoding of additions & corrections by the scribe and the recipient of the manuscript
 - Model only text, not non-textual elements (separating lines etc.)
 - Be as accurate & transparent as possible
 - Choices are product of individual interpretation
 - Consequences?
 - Create output for different research questions

<http://www.toufexis.info>

div1



div2 type= low



div2 type= high

Basic structure

- Columnar layout
- Thematoepistulae numbered by Crusius
- Page numbers provided by Crusius
- Vernacular in the left, learned in the right column

Encoding of abbreviations:

<abbr>&<expan>

ὄδ,

<expan>ὄ δεινα</expan>

πρῶτος
α ο

<abbr expan="πρῶτ">

α *from*

</abbr>ος

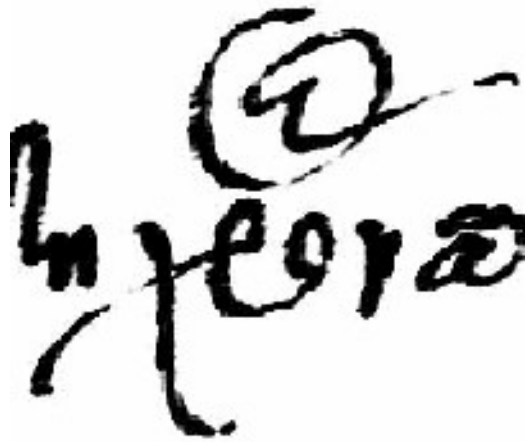
ἀρχή
α

ἀρχ<expan>ήν</expan>

<http://www.toufexis.info>

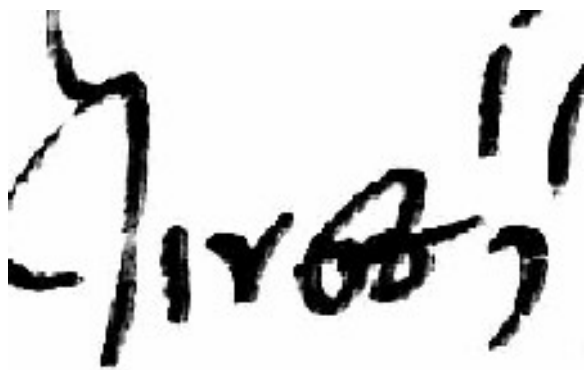
Letter combinations not encoded

Download



χρόνος

from



ΤΙΝΈΣ

<http://www.toufexis.info>

Alterations and corrections

Download

- `<add>` with attributes "hand" and "place" for additions by the scribe and the recipient
- `` with attributes "hand" and "type" for deletions by the scribe or the recipient
- `<unclear>` for cases that need verification

<http://www.toufexis.info>

Output for reading the text

Download

- HTML/CSS output
 - Keep the layout of the original manuscript as much as possible
- Make markup visible for human readers
 - Combination of colors, conventional markup (sigla, line numbering) & formatting (underline, line-through)
- Make this particular manuscript accessible to non-specialist readers
 - Allow them to build up their own mind about its peculiarities

<http://www.toufexis.info>

HTML output *Download*

- [diplomatic_transcription_v01.html](#)

from

<http://www.toufexis.info>

Next stage: Tokenization, segmentation

- Align vernacular and learned text at the level of sentence, phrase and word
 - Lemmatisation
 - Tokenisation
 - Segmentation in phrases and sentences
- Proceed with the diplomatic transcription or produce a normalized version of the text?

<http://www.toufexis.info>

Why produce a normalised version?

Download

- Avoid markup complexity
 - Nested elements:
 - `<w>ἀρχ<expan>ἦν</expan></w>`
- Information not necessary for a different level of linguistic analysis
 - Not a study of the graphematic system but of lexicon, morphology and syntax
- Can be linked with the diplomatic transcription
- Updating of changes more complicated

from

<http://www.toufexis.info>

Conclusions

Download

- Transcribing a manuscript electronically opens access to primary material for different groups of end-users
- Our choices at all levels affect the way our transcription will be used by end-users
 - It is very difficult to cover all needs
 - Make transcription available to others
- Digital transcription should be considered as a valid research tool for some kinds of primary material

<http://www.toufexis.info>