

Digital Italy
14th September 2021

Building a digital corpus and a computational lexicon of the languages of ancient Italy

Luca Rigobianco

Università Ca' Foscari Venezia



Università
Ca'Foscari
Venezia

Languages and cultures of ancient Italy Historical linguistics and digital models

Project of Relevant National Interest (PRIN)

<https://www.prin-italia-antica.unifi.it/>

- ❖ **Ca' Foscari University of Venice** (Department of Humanities)
- ❖ **University of Florence** (Department of Humanities)
- ❖ **National Research Council of Italy** (Institute for Computational Linguistics)

January 2020 – July 2023

Languages and cultures of ancient Italy

Historical linguistics and digital models

Objective

- ❖ **study of the languages and cultures of Ancient Italy**
(methods of epigraphy and historical linguistics; digital tools)

Expected results

- ❖ **digital corpus of inscriptions**
- ❖ **computational lexicon**

Focus on four languages: **Cisalpine Celtic, Venetic, Faliscan, and Oscan**

Digital corpus of inscriptions

- ❖ **edition of the text**
- ❖ **contextual information**
- ❖ **commentary**
- ❖ **translation**
- ❖ **bibliographical references**
- ❖ **photographs and drawings**

Digital corpus of inscriptions

❖ Identifier(s)	Vetter 245a, Giacomelli 4a, Bakkum 6
❖ Language	Faliscan
❖ Alphabet	Faliscan
❖ Typology	ownership inscription
❖ Object	cup
❖ Material	ceramic
❖ Ductus	dextroverse
❖ Origin	Falerii Veteres/Civita Castellana
❖ Date	6 th century BC
❖ Lemmata	Iartos
❖ Text Strings	art

Digital corpus of inscriptions

TEI-EpiDoc encoding

- ❖ interoperability
- ❖ availability
- ❖ sharing

Application to Restsprachen

- ❖ adaptability of XML
- ❖ bottom-up approach of TEI-EpiDoc

TEI-EpiDoc encoding Language and script

Identifier attribute

- ❖ **language subtag** (IANA Language Subtag Registry)
- ❖ **script subtag** (ISO 15924)
- ❖ **private use subtag**

```
<langUsage>
<language ident="grc-Latn">Ancient Greek written in Latin script</language>
<language ident="x-oldcam-latn-ci">Old Cam language in Old Cam script transliterated in Latin characters</language>
</langUsage>
```

TEI-EpiDoc encoding Language and script

```
<tei:scriptNote>
<tei:rs type="writingSystem" subtype="alphabet" ref="https://www.prin-italia-antica.unifi.it/p166.html#faliscan-faliscan">Faliscan alphabet</tei:rs>
<tei:rs type="wordDivision">scriptio continua</tei:rs>
<tei:rs type="syllabicPunctuationApplication"></tei:rs>
<tei:rs type="syllabicPunctuationSimplification"></tei:rs>
</tei:scriptNote>

<tei:langUsage>
<tei:language ident="xfa" source="https://iso639-3.sil.org/code/xfa">Faliscan</tei:language>
<tei:language ident="xfa-Ital-x-xfaxfa" source="https://www.prin-italia-antica.unifi.it/p166.html#faliscan-faliscan"
ana="https://unicode.org/iso15924/iso15924-codes.html">Faliscan in Faliscan alphabet</tei:language>
</tei:langUsage>
```

TEI-EpiDoc encoding Tokenisation

```
<tei:div type="edition" subtype="interpretative" xml:space="preserve">
<tei:div type="textpart" n="face_a" style="text-direction:r-to-l">
<tei:ab>
<tei:lb n="1" xml:id="Fal_6_1_1"/>
<tei:w xml:id="Fal_6_1_1_w_1">eko</tei:w> <tei:name type="praenomen" xml:id="Fal_6_1_1_w_2" ref="#p1">lartos</tei:w>
</tei:ab>
</tei:div>
</tei:div>
```

TEI-EpiDoc encoding DSL (Domain Specific Language)

- ❖ Readibility
- ❖ Compactness

HISTORY

#composition: @Falerii Veteres (@Civita Castellana) | 6th century BC

#finding: #discovered

@Falerii Veteres (@Civita Castellana) | La Penna necropolis, tomb LXXXVI/6

#by Adolfo Cozza, Angiolo Pasqui 1886

observation: #seen | Luca Rigobianco 2021

TEI-EpiDoc encoding DSL (Domain Specific Language)

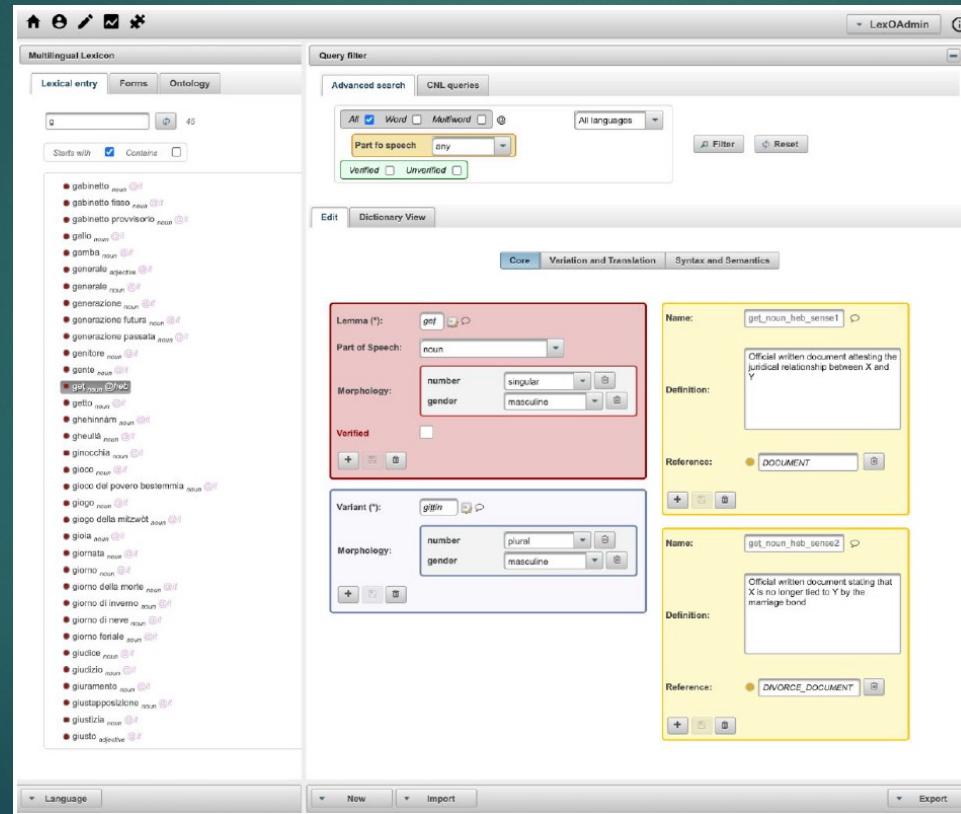
DSL → XSLT → XML

```
<tei:history>
<tei:origin>
<tei:origPlace type="composed">
<tei:placeName type="ancient" ref="https://pleiades.stoa.org/places/413126">Falerii Veteres</tei:placeName>
<tei:placeName type="modern" ref="https://www.geonames.org/6541123">Civita Castellana</tei:placeName>
</tei:origPlace>
<tei:origDate notBefore-custom="-0599" notAfter-custom="-0500" cert="medium">6th century BC</tei:origDate>
</tei:origin>
<tei:provenance type="found" subtype="discovered" when="1886">
<tei:placeName type="ancient" ref="https://pleiades.stoa.org/places/413126">Falerii Veteres</tei:placeName>
<tei:placeName type="modern" ref="https://www.geonames.org/6541123">Civita Castellana</tei:placeName>
<tei:placeName type="detailed">La Penna necropolis, tomb LXXXVI/6</tei:placeName>
<tei:persName>Adolfo Cozza</tei:persName>
<tei:persName>Angiolo Pasqui</tei:persName>
</tei:provenance>
<tei:provenance type="observed" subtype="seen" when="2021">
<tei:persName>Luca Rigobianco</tei:persName>
</tei:provenance>
</tei:history>
</tei:msDesc>
</tei:sourceDesc>
</tei:fileDesc>
```

TEI-EpiDoc encoding Linked Open Data

- ❖ **Geonames** (<http://www.geonames.org/>)
- ❖ **Pleiades** (<https://pleiades.stoa.org/>)
- ❖ **Trismegistos** (<https://www.trismegistos.org/>)
- ❖ **Getty Vocabularies** (<http://vocab.getty.edu/>)
- ❖ **Eagle Vocabularies** (<https://www.eagle-network.eu/resources/vocabularies/>)
- ❖ **Zotero** (<https://www.zotero.org/>)

Computational lexicon LexO



Computational lexicon

Epigraphy ID, word form, author, ...

Advanced search

Lexicon panel

Multilingual Lexicon

Epigraphies Lexical entry Forms Ontology

Search Starts with Contains

AL1 B01 BL1 ES121 ES127 PA1 PA2 PA4 PA6 PA20 PA21 PA24 PA26 PA28 PA43

PA2 Metadata Bibliography

Diplomatic Interpretative

plete.iwe.i.y.no
i.kara.n.mnio.i.e.
kupebari.s.e.xo

Pledei Veign o
i Karamni io i
ekupetari ego

Translation: I (am) the ekupetaris for Pledei- Venetic Karamnio-

xml

Edit Dictionary View

Lemma: **ekupetaris** n.m.
Stem: ekupetari-
StemType: io-stem
Suffix: *-ário-
Composition: endocentric compound
Compound members: *ekuo-; -pet-
Sense (certain):
Definition: funerary monument for a member of the ruling class
RelatedSense: *ekupetabos*
Translation: ekupetaris

Etymology:
IE-EtymologicalRoot: *h1lek-u-
Proto-It-EtymologicalRoot *ekwo-
Bibliographic Link [IEW, pp. 301-302, s.v. ekwo-s; EDL, pp. 192-193, s.v. equus](#)
IE-EtymologicalRoot *pot-
Proto-It-EtymologicalRoot *pot-
Bibliographic Link [IEW, p. 842, s.v. pot-is.](#)

Forms:
Variant1: **ekupetaris** n.m.s.Nom
Attested_in: Pa 2, lines 2-3 (certain)
Bibliographic link: [Pellegrini-Prosdocimi 1967, pp. 328-331](#)

Language New Import Export

Metadata panel

Epigraphy ID: PA 2
Object Type: stele
Material: trachyte
Dimensions: 86 h 66.5 w 34d
Writing system: inscribed (on the front side, along the right, upper, and left edges)
Alphabet: Venetic Subtype: Paduan type (syllabic punctuation properly applied)
Original Place: Patavium
Findspot: Padua, Casa Bassani
Date: 400-BC - 321BC
Epigraphic type: funerary inscription
Autopsy: n.a.

...

References panel

Bibliography:

G. B. Pellegrini, A. L. Prosdocimi, *La lingua veneta*, Padova 1967, I, 328-331

M. Lejeune, *Manuel de la langue vénète*, Heidelberg 1979

A. Marinetti, "Il 'signore del cavallo' e i riflessi istituzionali dei dati di lingua. Venetico ekupetaris", in *Produzione, merci e commerci in altino preromana e romana*, Roma 2003, 143-160

J. Pinault, "Venetic ekvopetaris and its Indo-European Background", *Wékkwos* 2016, 2, 179-194

Computational lexicon

Dictionary View

Lemma: *ekupetaris* n.m.

Stem ekupetari-

StemType: io-stem

Suffix: *-ārio-

Composition: endocentric compound

Compound members: *ekuo-; -pet-

Sense (certain):

Definition: funeral monument for a member of the ruling class

RelatedSense: *ekupetabos*

Translation: ekupetaris

Etymology:

IE-EtymologicalRoot: *h¹ek-u-

Proto-It-EtymologicalRoot *ekwo-

Bibliographic Link [IEW, pp. 301-302, s.v. ekuo-s; EDL, pp. 192-193, s.v. equus](#) ↗

IE-EtymologicalRoot *pot-

Proto-It-EtymologicalRoot *pot-

Bibliographic Link [IEW, p. 842, s.v. pot-is-](#) ↗

Forms:

Variant1: *ekupetaris* n.m.s.Nom

Attested_in: Pa 2, lines 2-3 (certain)

Bibliographic link: [Pellegrini-Prosdocimi 1967, pp. 328-331](#) ↗