

# ELEXIS travel grant - my visit at the Belgrade Center for Digital Humanities

Elina Boeva

I had the chance to spend the week of June 24-28 in Belgrade exploring the methods for retro-digitizing the *Latin-Bulgarian Dictionary* (ed. M. Voinov, A. Milev) currently in use in academic institutions. This was made possible by [ELEXIS - the European lexicographic infrastructure](#). As an ELEXIS grantholder, I got the opportunity to work with colleagues from the [Belgrade Center for Digital Humanities](#) and the [Institute for Serbian Language of the Serbian Academy of Arts and Sciences](#). The visit was supervised by [Toma Tasovac](#), the director of the BCDH, an ELEXIS partner and the leading Serbian institution exploring the use of data modelling, digital editions and standards-compliant lexicographic resources.

## 1. Travel to Belgrade

I travelled to Belgrade together with my colleagues Dr Dimitar Iliev (Assistant Professor in Ancient Greek and Latin at Sofia University, National Coordinator for [DARIAH-EU](#) in the framework of the [CLaDA-BG](#)) and Borislav Petrov, currently a BA student in Classics. The three of us form the core team of a future project for the retro-digitization of our Latin-Bulgarian dictionary, which we hope to be launched within the framework of the [CLaDA-BG National Infrastructure](#) in due course.

## 2. Institute for Serbian language

During the week, we worked in the magnificent building of the BCDH partner [Institute for Serbian Language](#) together with our Serbian colleagues and hosts - Toma Tasovac, Marija Gmitrovich, researcher in the Institute and BCDH affiliate, and Justina and Dimitrije, students of Classics at the University of Belgrade. Most of them already have serious experience in digitizing dictionaries. The BCDH and the Institute collaborated on various projects related to the Serbian lexicographic heritage, which lead to the creation of a dictionary platform <http://raskovnik.org/> and a platform for the transcription of handwritten heritage <http://www.prepis.org/>. We got a chance to explore both platforms, learn about the technologies behind them and consider what it would take to adopt similar approaches to creating our own online edition of the Latin-Bulgarian dictionary.



### 3. The dictionary - structure and inconsistencies.

**ob-dūco, dūxī, ductus** 3. 1. вода срещу, довеждам пред exercitum ad oppidum; posterum diem obducere прибавям. — 2. прокарвам (пред, около нщ.) fossam, vallum viis. — 3. а) навличам aliquid alicui rei; tene-

bras clarissimis rebus замъглявам, навличам мрак върху, покривам с мрак, затъмнявам; прен. callum dolori притъпявам чувствителността, ставам нечувствителен; пох или pubes caelo obducitur разстила се по небето, небето се покрива с. б) покривам arbores obducuntur cortice; frontem намръщвам; quae vastas obduxit което времето покри със забрава; frons obductus мрачно, намусено; пох obducta тъмна, мрачна; luctus, dolor obductus успокоила се болка. — 4. поглъщам, пия venenum.

**obdūco, duxi, ductum, 3. 1)** *йовесџи шџо на шџо* или *йроџив чеџа*: o. exercitum ad oppidum; o. Curium, поставити Курија као такмаца другим кандидатима. Отуда o. posterum diem, додати и други дан. 2) *нешџо љред нешџо љовуџи*: fossam ab utroque latere collis; o. seram, пустити резу. Отуда = *скуџиџи, сабраџи*: vela, увуџи; vestem; o. frontem, намрштити. 3) *нешџо љреко чеџа љревуџи*: o. tenebras rebus, помрачити; trop. o. callum dolori, отврднути према болу; obducta nocte, усред мрачне ноџи. 4) *љокриџи, обложиџи чим*: truncos cortice: alqd operimento; cicatrix obducta, рана зарасла. 5) *у се љовуџи* = *исџиџи*: venenum; potionem.

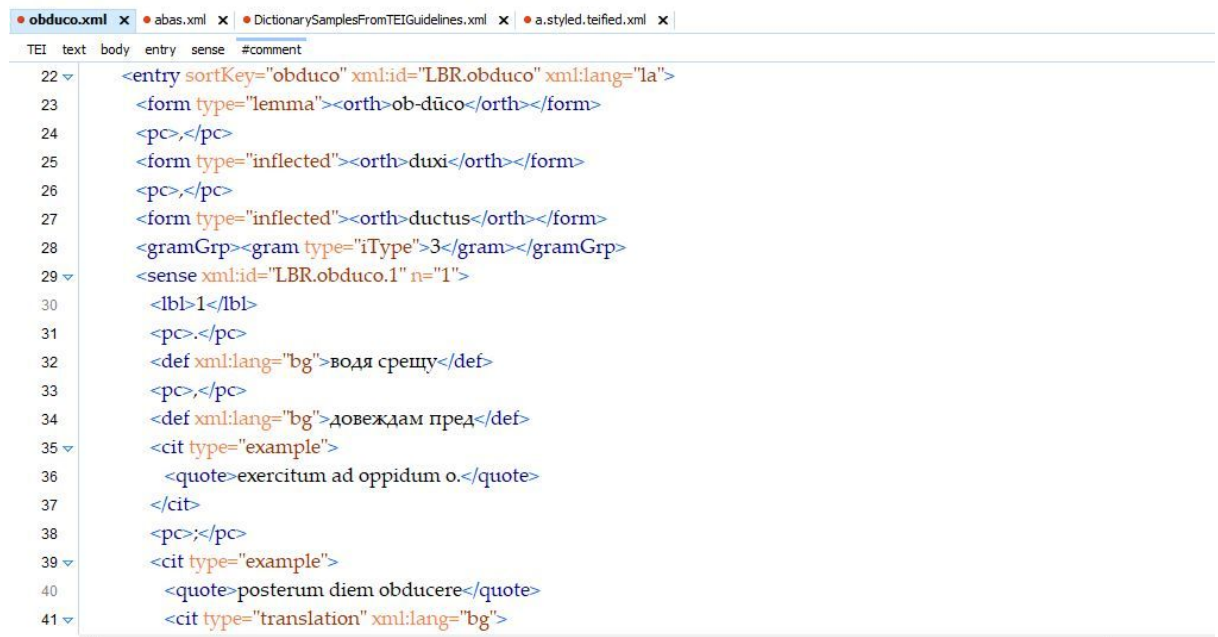
Our initial job was to perform a close analysis of the *Latin-Bulgarian Dictionary*, while our Serbian colleagues worked on a *Latin-Serbian Dictionary* (ed. Jovan Djordjevich). Both dictionaries may have been based on a German source,

possibly Wilhelm Freund's *Wörterbuch der Lateinischen Sprache*, which was, after all, also a major source for Lewis and Short as well, after Andrews translated it into English. The structure of the dictionaries is quite similar, but the editors from Bulgaria and from Serbia have made different decisions regarding, for instance, the use examples, abbreviations (e.g. in the Serbian dictionary the abbreviations are in Latin, while in ours they are mostly in Bulgarian) etc.

The first edition of the current Latin-Bulgarian dictionary was published in 1945. It was partially supplemented in the five subsequent re-editions. Although its structure is generally consistent, upon closer look, quite a lot of unexpected “peculiarities” showed up, which made me realise that the digitization is not the only job that is to be done - all the dictionary entries have to be revised and supplemented.

#### 4. Encoding

The next task was to encode our very first lemma under the guidance of Toma Tasovac, who introduced us to the TEI standards and tools, as well as the emerging TEI Lex-0 schema, which is currently under development also in the context of ELEXIS. He “walked” us through the elements, attributes and values used in lexicographic encoding in detail.



```
obduco.xml x abas.xml x DictionarySamplesFromTEIGuidelines.xml x a.styled.teified.xml x
TEI text body entry sense #comment
22 <entry sortKey="obduco" xml:id="LBR.obduco" xml:lang="la">
23 <form type="lemma"><orth>ob-dūco</orth></form>
24 <pc></pc>
25 <form type="inflected"><orth>duxi</orth></form>
26 <pc></pc>
27 <form type="inflected"><orth>ductus</orth></form>
28 <gramGrp><gram type="iType">3</gram></gramGrp>
29 <sense xml:id="LBR.obduco.1" n="1">
30 <lb>1</lb>
31 <pc></pc>
32 <def xml:lang="bg">водя срещу</def>
33 <pc></pc>
34 <def xml:lang="bg">довеждам пред</def>
35 <cit type="example">
36 <quote>exercitum ad oppidum o.</quote>
37 </cit>
38 <pc></pc>
39 <cit type="example">
40 <quote>posterum diem obducere</quote>
41 <cit type="translation" xml:lang="bg">
```

During the process of encoding, the need for supplements and changes in the dictionary entries became even more visible, so I decided to encode them all in XML and keep the original form as a comment.

```

157 <quote>luctus obductus</quote>
158 </cit>
159 <cit xml:lang="bg" type="translation">
160 <quote>успокоена скръб</quote>
161 </cit>
162 <p></p>
163 <cit type="example" xml:lang="bg">
164 <quote>dolor obductus</quote>
165 </cit>
166 <cit type="translation" xml:lang="bg">
167 <quote>успокоила се болка</quote>
168 </cit>
169 <p></p>
170 </sense>
171 </sense>
172 <!-- ORIG luctus, dolor obductus успокоила се болка -->
173 <sense n="4" xml:id="LBR.obduco.4">
174 <p>--</p>
175 <lbl>4.</lbl>
176 <def xml:lang="bg">поглъщам</def>

```

Thanks to Toma Tasovac’s detailed introduction, I got acquainted with the TEI guidelines for dictionaries and (as a sort of “homework”) was able to encode a few more entries and to explore the differences in encoding a verb, a noun, an adverb and an adjective. The overall approach we chose - with two teams working in parallel on their respective dictionaries while continuously updating each other and comparing each other’s work - turned out to be both fun and very productive.

## 5. Introducing hacks and tricks in Oxygen

One of the most useful things that we learned during our stay in Belgrade were the “hidden possibilities” of Oxygen. I had some experience using Oxygen in encoding epigraphic monuments, but the tips regarding shortcuts, code templates and styling options, for instance, which Toma introduced us to, were completely new to me, and they will greatly optimise not only the future encoding of the *Latin-Bulgarian Dictionary*, but my work in digitization in general.

## 6. XPath & XSLT

The real challenge was to learn using XPath to navigate through xml-encoded dictionaries. This, thankfully, was made easier by numerous real-life examples we looked at and the exercises that Toma prepared for us, which finally made me feel confident enough to write expressions on my own - and they worked! We were also introduced to the application of some basic XSLT transformation scenarios, that would make our work on the digital Latin-Bulgarian dictionary a lot easier.



## 7. Cultural program

Though our program was intense and I was completely focused on all the new information, I still managed to appreciate the beauty of Belgrade and visit some of the sights and museums. Belgrade is an unique city and I enjoyed every moment of my stay. We had lots of fun with Toma, Marija, Justina and Nikola (especially laughing at the sound of each-other's language).

The overall experience with the [Belgrade Center for Digital Humanities](#) made me determined to continue and extend my work on the digitization of the Latin-Bulgarian dictionary. I gained an enormous amount of knowledge, that would make the start of the dictionary project much easier. I sincerely hope that we'll continue working closely with our Serbian colleagues and that their experience will help us in our initial steps towards the edition of the much needed digital Latin-Bulgarian dictionary.