

Comprehensive Coptic Lexicon V 1.0



Release Notes

May 15, 2019

Publishers:	Database and Dictionary of Greek Loanwords in Coptic (FU Berlin) Strukturen und Transformationen des Wortschatzes der ägyptischen Sprache (BBAW)
Distributor	Freie Universität Berlin/Refubium
Raw data:	http://dx.doi.org/10.17169/refubium-2333
Processed data:	http://coptic-dictionary.org
Compatibility:	TEI Dictionary Module, TLA XML Schema
Availability:	CC BY-SA 4.0
Technical specification:	1. DDGLC_TLA_Conversion_Procedure.pdf 2. DDGLC_CLL-TLA_Concordance.pdf 3. Form_treatment3.pdf 4. Metadata_Refubium_TEIHeader.pdf

1. Description:

Version 1 of the Comprehensive Coptic Lexicon is the joint release of the “Database and Dictionary of Greek Loanwords in Coptic” (Freie Universität Berlin) and the research project “Strukturen und Transformationen des Wortschatzes der ägyptischen Sprache” Thesaurus Linguae Aegyptiae” (Berlin-Brandenburgische Akademie der Wissenschaften). It includes the following features:

- Version 3 of the BBAW lexicon of Coptic Egyptian with a number of issues resolved (see §3. Resolved bugs).
- Version 1 of the DDGLC lexicon of Greek loan words in Coptic.
- New TLA TEI XML headers.
- Extended TLA TEI XML Schema.

The DDGLC serialized “raw” data have been imported and compiled according to TLA XML Schema to include **3229** new entries and **10029** new forms (standard forms, i.e. <form type=“lemma“, are excluded from this number as they are copies of the existing non-lemma forms). The ETL procedure extended over the following milestones:

December 2017:	Initial DDGLC database export as XSLX
June 2018:	Clean-up and conversion of XSLX to TLA XML
October 2018:	Export of TLA XML as XSLX for debugging and proofreading by DDGLC
February 2019:	Conversion of the proofread XSLX file to TLA XML
April 2019:	Conceptualization of the new TLA TEI XML headers Adaptation of TLA XML Schema
May 2019:	Final proofreading and production of the Release Candidate 1.

BBAW and DDGLC datasets were compiled into a single file, entitled “Comprehensive Coptic Lexicon”, containing now **11272** entries and **20141** forms. For detailed information concerning the ETL procedure please refer to the technical specification. For referencing the raw data, please use the Refubium DOI (<http://dx.doi.org/10.17169/refubium-2333>), for referencing the processed data please use the publication: <http://aclweb.org/anthology/papers/W/W18/W18-4502>.

2. Known issues

- 113 Greek-Coptic forms have no dialect information assigned ("?" used instead) because they are not yet attested in the DDGLC corpus but are required to resemble the standard form of the lemma.
- Nouns deriving from Greek roots differentiated in gender (e.g. ἀγγελικη (fem. in Geek), ἀγγελικος (masc. in Greek), ἀγγελικον (fem. in Greek)), are treated as separate lemmata, but encode no gender information in Coptic.
- Verbs originating from Greek roots through different borrowing mechanisms (e.g. (εϣ-) + ἀγαπαν, Β (aux + Greek infinitive) vs. ἀγαπα, S (Greek imperative); ἀναχωριν (Β, Greek infinitive) vs. ἀναχωρει (S, Greek imperative)) are treated as separate lemmata.
- “Senses” in Greek-Coptic dataset provide English translation only.

3. Resolved bugs

#	Bug (https://github.com/KELLIA/dictionary/issues)	Reported by
#34	Disambiguate C324, C370, C395, C404 containing two parts of speech in the same entry (Vb. and Adj.)	Kupreyev
#38	Minor errors (reported by F. Hoffmann)	Hoffmann
#39	ζοογ should be VSTAT	Zeldes
#40	Übersetzung "Kürbis" in der koptischen Wortliste nicht korrekt	Popko
#41	Tooltips for bibliography	Feder
#42	Correct CF14076 (μντωεερε ωνμ -> μντωηρε ωνμ)	Miyagawa
#46	Add IDs to the new entries (see #34)	Kupreyev
#47	Lemma for aueis	Zeldes
#48	Missing forms of εζβηητε 'foam'	Zeldes

TLA expresses gratitude to F. Hoffman, S. Miyagawa and L. Popko for reporting the new bugs.

3. Files to be released:

File name	File size (bytes)	Modification date
BBAW_Lexicon_of_Coptic_Egyptian-v3-2019.xml	7,871,351	15.05.2019 7:46
Comprehensive_Coptic_Lexicon-v1-2019.xml	13,200,133	15.05.2019 7:47
DDGLC_Lexicon_of_Greek_Loanwords_in_Coptic-v1-2019.xml	4,611,799	15.05.2019 7:47
Coptic_Lemma_Schema-v1.1.xsd	27.627	14.05.2019 13:28