Building an interactive Bohairic Coptic thematic lexicon

Digital Coptic III, 12/7/2020

Ἁλέλη Ἁρέλ (Halely Harel)
Πολις (Polis, The Jerusalem Institute of Languages and Humanities)

PhD Candidate, Hebrew University of Jerusalem
Mandel School For Advanced Studies in the Humanities Doctoral Fellow

haleli.harel@mail.huji.ac.il
Objectives of the project, working title, ‘ⲥⲱⲧⲉⲙ’

**RESOURCES:**
to create a thematic lexicon that documents examples of each lexical item- Which genre, which time.
(and by that) Gradually creating a linked and annotated corpora of Bohairic.

**COMMUNITY:**
To offer a space which facilitates a digital publication of teaching materials and promotes an open discussion about the lexicon (and grammar, phonology) of “spoken” Bohairic.

ⲥⲱⲧⲉⲙ is not online yet, it will soon be available at:
https://www.polisjerusalem.org/language/coptic/,
https://sotem.info
Creating resources for teaching Coptic in full immersion

Project publications
(Polis Institute Press, 2021):

1) Ⲟⲛⲧⲣⲟ- A textbook for spoken Bohairic in 16 lessons.

2) Ⲟⲟⲩⲓ, Ⲣⲧⲧⲓ, Ⲥⲟⲩⲓ- A practice book in spoken Bohairic.
Teaching with TPR and Living Sequential Expression
Teaching with TPR and Living Sequential Expression
Building an interactive Bohairic Coptic thematic lexicon
Building an interactive Bohairic Coptic thematic lexicon
Q: What does BOHAIRIC stands for?
A: When we say Bohairic Coptic we refer to a several linguistic varieties:

WRITTEN (LITERARY) EVIDENCE

Early varieties- earlier sources, paleo/old Bohairic.
(Such as P. Bodmer III; P. Vat. Copto 9; P. Heid. Kopt 452; P. Mich. Inv. No. 926; P.Mich.Copt. III.3; BL Or. 739; Bal. 19.)
(aka “B4” and it sub-varieties aka B74 (Type of south Bohairic), B74!, and B504-507*?.)

-no written evidence from 5th to 9th centuries.-

“Nitrian” (=Scetian)- texts originating from the Monastery of St. Macarius.

‘Scriptural’/ classical bible “medieval” Bohairic- (aka “B5”)
Q: What does BOHAIRIC stands for?
A: When we say Bohairic Coptic we refer to a several linguistic varieties:

**SPOKEN EVIDENCE**

1/ An Egyptian dialect of the region of lower Egypt, gradually declining as a spoken language after Arab conquest.

2/ Bohairic as a liturgical language of entire Coptic church - Early 2nd millennium, up to this day.

3/ Various attempts of spoken Bohairic in last two centuries. Based on traditions, scriptures, liturgy and on some lexical innovations.
Building an interactive Bohairic Coptic thematic lexicon

So what is our actual Corpus?

- Biblical manuscripts. The core corpus, OT and NT.
  
  - OT
    - Vat. copto 1, Paris copte 1? (ASH, Topics 2007).
  
  - NT
  
  - Job, Psalms, Proverbs
  
  - Major Prophets
  
  - Patristic texts
  
  - Martyrologies
  
  - Homelies
  
  - Liturgical compositions
  
  - Philological works
  
  - Private inscriptions (and Graffiti) (Such as documentary texts at Kellia).
Building an interactive **Bohairic Coptic** thematic lexicon

- Starting by documenting selected OT and NT medieval Bohairic source manuscripts.
- Aiming to create corpora by standards offered by Coptology projects.
Building an interactive Bohairic Coptic thematic lexicon

Grammatical progression-
Identifying literary and spoken Coptic varieties.

Introduce grammar gradually, accept traditional “conversational” forms that are not documented in written language, but aim to instruct literary grammar for its features and complexity.

If your right eye causes you to stumble, gouge it out and throw it away. It is better for you to lose one part of your body than for your whole body to be thrown into hell.
(Matthew 5:29)
A practice of grammatical theory

The choice of texts in our textbook is based also on literary varieties of speech representation.

- Recording as much typical direct speech grammatical forms. (Richter 2006).

- using specific narrative vs. dialogue forms (Shisha-Halevy 2007), for telling a story.
- Roots of western Coptic lexicography in “Sullam”/”Scala” type works.

Here, Kircher’s publication Of Shams al-Riyasah al-Quss ibn Kabar (d. 1324)’s *Al-Sullam al-Kabir*. 

Copto-Arabic philological works and grammars are the starting point of our project.
Building an interactive Bohairic Coptic thematic lexicon

Champollion’s unpublished grammatical notes

https://gallica.bnf.fr/ark:/12148/btv1b10085358n/f28.image
Building an interactive Bohairic Coptic thematic lexicon

- Citing Labib and others for innovations and traditions that are not part of standard lexicography.
La Croze, M. V. and C. G. Woide. 1775. 
Lexicon Aegyptiac-Latinum. 
Lacroze-Scholz-Woide.

Peyron, Amedeo. 1835. 
Lexicon linguae Copticae.
Building an interactive Bohairic Coptic thematic lexicon

<table>
<thead>
<tr>
<th>C3972</th>
<th></th>
</tr>
</thead>
</table>

**Form**
- Dial.: S
- ID: CF5884
- POS: Vb
- Attestation: C

**Scriptorium tag:** V

1. (En) write
   - (Fr) écrire
   - (De) schreiben
   - **Bibliography:** [CD 881](#), CED 17:1673, KoptHWb 21:1, DELC 20:1, ChLCS 59

2. (En) register
   - (Fr) enregistrer
   - (De) eintragen, verzeichnen
   - **Bibliography:** [CD 881](#), KoptHWb 21:1, ChLCS 59

3. (En) draw, paint
   - (Fr) dessiner, peindre
   - (De) malen, zeichnen
   - **Bibliography:** [CD 881](#), CED 17:1673, KoptHWb 21:1, DELC 20:1, ChLCS 59

**See also:**
- Σεις (σης) – ascribe to, inscribe, write down
- Σεις (σης) – write outward, publish
- Σεις (σης) – write in (to place), inscribe, register
- Σείς – writing, letter (epistle or alphabetical)
- Νιστέρα – writing
- Κτέσις – without letters, illiterate
- Ροκέτα – writer, scribe
- Σκοιτά – act, form of writing
- Σέγ – 1. writer, teacher, master
- Σέγ – 2. master (of a craft)
- Σή – written copy, diploma
- Νιστέρα – skill
- Ροκέτα – become, be teacher, skilled
A lexical item includes the following details:

- **Bo-ID** - form and item no. in the ṭⲧⲉⲛⲥⲁ ϫ ϭ ₪ dataset. χⲇⲉ 40/381b

- **TLA-ID** - TLA lemma no. C3972 *(may have several ids)*

- (dialectal) **Form-ID of CDO** - not available in this case. *(may have several ids)*

- **Other dictionaries ids of CDO** - e.g. CD 381b⁷; CED 172⁷-173; KoptHWb 21l⁷; DELC 204⁷; ChLCS 53b⁷

- Translations (from CDO, EN, DE, FR) + Arabic translation and its source + meaning in Coptic (possible).

- Themes, entities tagging.

- marking neologisms. Linking them to form/TLA-IDs
Building an interactive Bohairic Coptic thematic lexicon

**Tokens**- examples from source texts.
- context line in unicode.
- glossed TEI format.
- Links to witnesses (which are annotated for their metadata, CMCL ids etc.)

I.e.  الصحيح، إعداد الحزام وأخذه ǎnd ǎnd ǎnd ǎnd أخذهم ǎnd أخذهم ǎnd أخذهم ǎnd أخذهم.

‘Saying, Master, Moses wrote unto us, If any man's brother die, having a wife, and he die without children, that his brother should
take his wife, and raise up seed unto his brother.’

An example includes a full “sentence,” with POS tagging and glossing- Using the Coptic

Sscriptorium NLP service.

- Commentary of author. Links to additional relevant publications.
- Sound recording/s.
- (In the UI)- Users comments.
# Tagging ‘themes’

- Going over the Bohairic portion of Crum, tagging lexical fields [https://concepticon.clld.org/](https://concepticon.clld.org/) and themes, (entities?).
- Marking themes/categories suggested by medieval philological works.
- Themes, following the Polis method are based on Modern world experience and daily life (kitchen, bathroom, school, street etc.)

#Planned thematic “open calls” for specialized language (e.g. medical terms, flora and fauna)
Users/Researchers can contribute:
- a reference, comment, opinion.
- a token (new example of a lexical item).
- Suggestion for a new lexical item.
- audios and video of the course songs and texts.

This is a collaborative project!
Which means it will be reviewed and later co-authored by many authors-

Guidelines will be available in sotem.org
Building an interactive Bohairic Coptic thematic lexicon

Within the first stage of the project (2/2021), website will be up with the database of all words in the coursebook. (not online yet, will be on https://www.polisjerusalem.org/language/coptic/, https://sotem.info)

Our goal: Making traditions of lexicography and textual studies more accessible to students of spoken Coptic.

Using building blocks that are the tools presented by Coptic DH projects, all data will be reviewed and published for reuse by other projects.

Gradually, the structure of the lexicon, frequency, hapax cases, and loans will be evaluated and displayed.
Τῷ εὐαγγελίῳ τοῦ Ἰωάννη (καὶ ἀρετευματωμένος ἐρωτ)

"Τῷ εὐαγγελίῳ τοῦ Ἰωάννη (καὶ ἀρετευματωμένος ἐρωτ)
(Ἰωάννης Ιωαννίνην, Ιωαννίνην)"
References


Peust, C. 1999. Egyptian Phonology, Monographien zur Agyptischen Sprache 2, Goettingen


