glottothèque
languages of Anatolia, Caucasus, Iran, Mesopotamia

Armenian
Tools & Resources

Victoria Khurshudyan

Inalco, SeDyL, CNRS, IRD

Paris, January 24, 2022

visit glottothèque at: https://spw.uni-goettingen.de/projects/lacim/

Classical Armenian 1

Corpora & textbases	**Classical Armenian Bible** with parallel King James Version, Arak29 foundation · Bible text with full morphological annotation · aligned corpus (CA – Eng) · 630.000 tokens (60.000 unique tokens, 12.000 lexemes)
	GREgORI project (UCLouvain) · Hellenophile Classical Armenian texts (6th-7th · cen.), · translated texts from Greek · 66.812 tokens (16.000 unique tokens) with full morphological annotation B. Coulie, B. Kindt, G. Kepeklian (UCLouvain), Digital Armenian International Conference (Inalco, 2019)
	Digilib (American University of Armenia) · comprehensive text database (5th - 18th)
	TITUS project (Johann Wolfgang Goethe University) · Full Bible text aligned with the Greek version
	Leiden Armenian Lexical Textbase (University of Leiden) · Textbase (Bible, classic native works, · lexical annotation, POS tagging · aligned corpus of Gospel of John and the Book of Jonah (CA-Greek). · corpus search functionalities

Classical Armenian 2

Dictionaries	**Calfa** - comprehensive online reference dictionary platform
	· 54.000 headwords with 150.000 examples (1.3 million tokens, 190.000 unique tokens)
	· Awetikʻean Gabriēl, Siwrmēlean Xačʻatur, Awgerean Mkrtičʻ, New dictionary of the Armenian language, Venise, 1836-1837.
	· Bedrossian Matthias, New Dictionary Armenian-English, Venise, 1875.
	· Calfa Ambroise, Dictionnaire arménien-français, Paris, 1861.
	· J̌axǰaxean Manuēl, Dizionario armeno-italiano, Venise, 1837.
	Leiden Armenian Lexical Textbase (University of Leiden)
	· New dictionary of the Armenian language, Venice, 1836
	· Ačaṙyan Hratchya, Armenian Root Dictionary, Yerevan, 1926-35
	· Matthias Bedrossian's New Dictionary Armenian - English, Venice, 1879
	· Baṙgirk' Hayoc', (1698) 1975, Erevan

Eastern Armenian

Corpora & textbases	**Eastern Armenian National Corpus** · a comprehensive corpus with about 110 million tokens · a powerful search engine for making complex lexical morphological queries · a learner's corpus including English translations for frequent tokens · a diachronic corpus covering SEA texts from the mid-19th century to the present · a mixed corpus consisting of both written discourse and oral discourse · an annotated corpus with morphological and metatext tagging · an open access corpus · an electronic library with full access to over 100 Armenian classic titles V. Khurshudyan, International Conference Digital Armenian (Inalco, 2019)
	Universal Dependencies treebank · 2.502 manually annotated sentences (about 53.000 tokens [v. 2.5]) · full morphological and syntactic annotations
	Russian National Corpus · aligned corpus (EA – Russian) · 2,4 mln tokens · full morphological annotation
Tools	**EANC's source code**

Western Armenian

Corpora & textbases	**Universal Dependencies treebank** · 2.524 manually annotated sentences (about 40.000 tokens [v. 2.8]) · full morphological and syntactic annotations
	Digilib (American University of Armenia) about 3 mln tokens (19th - 20th)
Dictionaries	**Parastan.fr** thematic multi-entry French-Western Armenian Dictionary *A. Donabedian, International Conference Digital Armenian (Inalco, 2019)*
Tools	**Boyacioglu**'s verb paradigms and lexicon Boyacioglu, Nisan. *Hay-pay: les verbes de l'arménien occidental guide pratique de conjugaison*. Paris: l'Asiathèque, 2010.

common resources

Nayiri: the biggest library of electronic Armenian dictionaries with 132 electronic dictionaries of different types

Transliteration tool
(for an overview on transliteration systems for Armenian see here)

Wiktionary
- 15 000 headwords on English wiktionary
- 303 005 headwords on Armenian Wiktionary
- word definition
- etymology
- inflected forms (declension, conjugation)
- pronunciation (only on English Wiktionary)

Fundamental Scientific Library: the biggest Armenian bibliographic database with considerable digitized resources

Grahavak: Comprehensive thematic catalogue of Armenian online resources

this lecture

is part of the series *Glottothèque: Languages of the Anatolia, Caucasus, Iran, Mesopotamia; grammatical snippets online*, ed. by. C. Bulut, A. Donabédian-Demopoulos, G. Haig, G. Khan, P. Samvelian, S. Skopeteas, N. Sumbatova. Bamberg/Cambridge/Göttingen/Moskow/Nicosia/Paris: LACIM network.

You may find related lectures and further information at the Glottothèque website at: https://spw.uni-goettingen.de/projects/lacim/

LACIM

glottothèque

Anatolia, Caucasus, Iran, Mesopotamia
grammatical snippets online