**Generalities**

The database aims at creating a comprehensive lexicon of the Modern South Arabian language Soqotri, the mother tongue of more than 100.000 inhabitants of the island of Soqotra in the Arabian Sea.

In spite of its immensely archaic nature and the exceedingly rich oral tradition of the islanders, Soqotri is still poorly known to the Western scholarly audience and practically unknown to academics and laymen of the Arab and Islamic world. Lexical resources of Soqotri are copious widely beyond one’s imagination and, as such, can be legitimately compared to the proverbially rich thesaurus of the pre-Islamic and early Islamic Classical Arabic. A fundamental account of Soqotri vocabulary is, therefore, more than a challenging task.

Vitaly Naumkin’s first trips to the island took place in 1970–1980s, when a fundamental understanding of the complicated rules of Soqotri grammar could be achieved and a first approximation to the oral literature of the islanders was made. Since 2010, Naumkin, Bulakh and Kogan carry out a still broader enterprise of documenting, analyzing and explaining to the public the traditional lore of Soqotra and the deeply original portrait of its unique tongue. More than 30 articles dedicated to *Soqotrana* have been published meanwhile, as well as two volumes of our *Corpus of Soqotri Oral Literature* (CSOL I–II) which appeared in 2014 and 2018 in the renowned Brill publication house (Leiden, the Netherlands).

The lexicon aims to be as comprehensive as possible and will take into account the following major sources for the Soqotri vocabulary known so far.

The glossaries of the two published volumes of our *Corpus of Soqotri Oral Literature* (CSOL I–II) and the third volume (CSOL III), in an advanced stage of preparation.

The lexical material of the earliest extensive publications in Soqotri, viz. the records of the Austrian South Arabian Expedition (1902–1905–1907), published and annotated by the great Austrian orientalist D.H. Müller.

The vocabulary contained in the forthcoming collection of Soqotri traditional poetry collected and translated by the Scottish researcher Miranda Morris. For many years, our team maintains an excellent working relationship with Dr. Morris and can rely on the advice from this leading authority of today’s Modern South Arabian studies.

The forthcoming Lexicon pursues two different, yet closely related goals.

On the one hand, it is conceived as as fully-fledged scholarly publication, to be produced by the renowned Brill publication house in the Netherlands. This prestigious publication will certainly become the main reference tool for the study of the Soqotri language and, more broadly, the oral traditions of the Arabian civilization.

On the other hand, it will play a key role in establishing Soqotri as an efficient written medium on the island, which will be achieved through a systematic application of the Arabic-based Soqotri writing developed and successfully implemented by the Soqotri-Russian working team. Moreover, all Soqotri words in the Lexicon will be translated not only into English, but also to Standard Literary Arabic, which will make its riches fully accessible to the Arabic-reading public in the UAE, the Gulf region and throughout the Arab and Islamic world. Last but not least, it will give a serious impetus for comparable initiatives pertaining to Soqotri’s close sister tongue Mehri (the Mahra province).

**Participants**

The Russian research group, headed and directed by Vitaly Naumkin (Institute for Oriental Studies, Russian Academy of Sciences) comprises three senior researchers, each primarily responsible for a certain scopes of creative and technical tasks. Ilya Arkhipov (Institute for Oriental and Classical Studies, HSE University) has developed and will maintain the technical side of the database, being also involved in the Semitic etymological research tasks and the preparation of the audio materials for individual lexical items. Maria Bulakh directs the integration of the existing lexical materials into the database’s framework. Leonid Kogan functions as the general coordinator of the project and is responsible for Semitic, MSA and Arabic etymological glosses.

**Scope and sources**

All lexical items featuring in the database are expected to meet two fundamental criteria: to proceed from **published** sources and to be **verified**, in form and meaning, by Soqotri native speakers. Both criteria eventually aim at a single goal, viz. maximum reliability of the lexical evidence presented.

Restriction to published data provides the user with an efficient means of control, especially in the semantic domain: in almost every case, “published” means here “published in context”.

Verification by native speakers means that, at least in the speech of the Da‘arho language community, a given word has a certain well-defined consonantal and especially vocalic shape, assured not only acoustically, but also in writing. It also assures an authoritative judgement on the basic and peripheral meanings of the lexeme, in most cases shared by at least three native speakers. There is no need to stress that other speakers of Soqotri, belonging to different geographical areas, age categories etc. will be able to polish and refine these phonological and semantic facts. This is a key direction for further research of our team, yet we are strongly convinced that the nuclear part of any serious lexicographic description is to be dialect-bound and corpus-bound.

**Sources presently included**

● Catalogues of verbal lexemes appended to summary descriptions of the verbal morphology of Soqotri and text illustrations therein

V. Naumkin, M. Bulakh, D. Cherkashin, L. Kogan, A. Issa, I. Gumaan. Studies in the Verbal Morphology of Soqotri I/2. Strong Triconsonantal Roots in the Basic Stem (the Lexical Data). *ZAL* 60 (2014):35–73.

V. Naumkin, M. Bulakh, D. Cherkashin, L. Kogan, A. Issa, I. Gumaan, M. Mohammed. Studies in the Verbal Morphology of Soqotri II: Weak and Geminated Roots in the Basic Stem. *ZAL* 63 (2016):19–60.

V. Naumkin, M. Bulakh, D. Cherkashin, L. Kogan, A. Issa, I. Gumaan, M. Mohammed. Studies in the Verbal Morphology of Soqotri III/1: the Second Stem. *ZAL* 69 (2019):61–93.

V. Naumkin, M. Bulakh, D. Cherkashin, L. Kogan, A. Issa, I. Gumaan, M. Mohammed. Studies in the Verbal Morphology of Soqotri III/2: a List of Sound and Weak Verbs Belonging to Second Stem in Soqotri. *ZAL* 70 (2019):73–91.

M. Bulakh, L. Kogan, A. Issa, I. Gumaan, M. Mohammed. The Causative Stem in Soqotri. The Analysis. *BJALL* 12 (2020):260­–285.

M. Bulakh, L. Kogan, A. Issa, I. Gumaan, M. Mohammed. The Causative Stem in Soqotri. The Data. *BJALL* 13 (2021):239–28.

M. Bulakh. The Semantics and Syntax of Stem IV in Soqotri. JSS 66 (2021):263–292.

● The “Soqotri Lexical Archive”

V. Naumkin, L. Kogan, D. Cherkashin, A. Issa, I. Gumaan. Soqotri Lexical Archive: the 2010 Fieldwork Season. *ZDMG* 163 (2013):61–85.

V. Naumkin, L. Kogan, D. Cherkashin, A. Issa, I. Gumaan. Soqotri Lexical Archive: the 2011 Fieldwork Season, *ZDMG* 165 (2015):44–61.

V. Naumkin, L. Kogan, D. Cherkashin, A. Issa, I. Gumaan. Soqotri Lexical Archive: the 2012 Fieldwork Season. *ZDMG* 166 (2016):57–80.

● Miscellanneous text publications of the Russian-Soqotri team not included into CSOL

V. Naumkin, M. Bulakh, L. Kogan. Two Erotic Stories from Soqotra Revisited. *Babel und Bibel* 7 (2014):527–563.

V. Naumkin, L. Kogan, I. Gumaan, A. Issa, D. Cherkashin. *Soqotri Texts in the Phonogrammarchiv of the Austrian Academy of Sciences. An Annotated Edition*. Oxford, 2015.

I. Gumaan, L. Kogan, D. Cherkashin. “The Lord”: An Apology for the Muslim Faith from the Island of Soqotra. *JSS* 64 (2019):535–563.

L. Kogan. Сокотрийские колыбельные. *Orientalistica* 3 (2020):443–456 (<https://doi.org/10.31696/2618-7043-2020-3-2-443-456>).

В.В. Наумкин, Л.Е. Коган. Сокотрийская народная поэзия: вечное возвращение. *Антропология и этнология*: *современный взгляд*. Moscow, 2021. Pp. 522–534.

**Sources to be included (2022–2024)**

● The glossary of CSOL III, scheduled for publication by the beginning of 2023.

● A few forthcoming publications by the Russian-Soqotri research team, such as the next issue of the “Lexical Archive” by Naumkin, Kogan et al. (scheduled for 2022) or Bulakh’s booklet on the passive-reflexive (VIII) stem in Soqotri (scheduled for 2023).

● Lexical materials of the “Vienna Corpus” (SAE IV, VI and VII), achieved through a systematic perusal of Leslau’s *Lexique Soqotri* and an in-depth inquiry into the pertinent passages from the SAE texts.

● Scores of “new” lexemes scattered over *The Oral Art of Soqoṭra* by Miranda Morris and Ṭānuf Sālim Nuḥ Di-Kišin (Leiden–Boston, 2021) and featuring in the annotated glossaries therein.

● Botanical terminology of *Ethnoflora of the Soqotra Archipelago* by A. Miller and M. Morris (Edinburgh, 2004). The book, among its many merits, is a valuable source of various categories of non-botanical terminology (landscape and meteorology, foodstuffs, building and household terms). Besides – and quite importantly – many (if not most) Soqotri plant names are, in fact, derived from common nouns and adjectives, often missing from our files.

● Animal names from *Fauna of the Socotra archipelago* by W. Wranik and Omar Al-Saghier (Rostock, 2003).

● A wealth of anatomic terminology found in the four articles by M.-C. Simeone-Senelle and A. Lonnet (and especially 1991). A preliminary impression is that “new words” missing from our files as well as from Leslau’s *Lexique* are quite numerous.

● Nakano, Aki'o. 1986. *Comparative vocabulary of Southern Arabic: Mahri, Gibbali, and Soqotri*. Tokyo: Institute for the Study of Languages and Cultures of Asia and Africa, Tokyo University of Foreign Studies.

● ранние публикации В.В.Наумкина

All lexemes known to our language consulants will be verified by them for form and meaning.

**Visual materials**

**Sound materials**

**Etymology**

In spite of being primarily synchronical in its scope, the dictionary will be sensitive to etymological issues, in an attempt to attract the attention of comparative Semitists to the invaluable lexical treasuries of Soqotri (and Modern South Arabian in general). Besides, in certain cases Semitic etymology can be helpful in eliciting the semantic nuances of rare and obscure Soqotri words and the diachronic path of their semantic evolution.

**Other MSA**

**Semitic in general**

**Dialectal Arabic**

Constant attention to Arabic loanwords is among the hallmarks of the project, which presupposes a systematic perusal of the extant dialectal vocabulary of the South Arabian Arabic varieties. One can only regret that many of them are so imperfectly (if at all) described from the lexicographic point of view, particularly those most naturally expected to be the main sources of the Arabic borrowings on the Island (such as Adeni, Hadrami or Dhofari).