

Towards the documentation of Shughni language: building on online dictionary and corpus

Shughni is a Pamir language spoken in mountainous areas of Tajikistan and Afghanistan by approximately 80 000 speakers (according to Ethnologue 2016). There are two Shughni-Russian dictionaries published in Russia: one was compiled by Ivan Zarubin (Zarubin 1960) in the first half of the 20th century, the other by Dodkhudo Karamshoev in the second half of 20th century (Karamshoev 1988). Besides, (Zarubin 1960) includes a collection of spontaneous texts.

This paper presents a part of a bigger project of Shughni documentation and is aimed at systematizing of the available linguistic data through designing two online tools a dictionary and a corpus.

The dictionary is based on both previously mentioned sources, so that the user of the resource can choose between them or observe entries from both sources simultaneously and follow the evolution of meanings during the 20th century. The large majority of dictionary entries is supplied with morphological notation (i.e. inflection classes), paradigmatic and derivational information and annotated examples. For some words dialectal variants are provided. The online dictionary provides a reverse translation (to Shughni) from two languages: Russian and English.

The second part of the resource presented is a morphologically annotated corpus, based on texts from (Zarubin 1960), fragments of the Gospel of Luke and several oral texts collected from a native Shughni speaker.

Two transcription systems have been used for Shughni: one based on the Cyrillic script and one based on the Latin script. Both the corpus and the dictionary are supplied with a transliteration module which provides search in both scripts.

These tools could make Shughni materials available not only to Russian speaking scholars and prove useful for a wide range of linguistic studies as will be shown in our presentation.

Sources:

Karamshoev D. (1988) Shughni-Russian dictionary in 3 volumes. Moscow.

Karamshoev D. (2005) Russian-Shughni dictionary (Fourth Volume).

Ed.: Steblin-Kamenskiy I.M. Dushanbe.

Lewis M. Paul, Gary F. Simons, and Charles D. Fennig (2016) Ethnologue: Languages of the World, Nineteenth edition. Dallas, Texas: SIL International.

<http://www.ethnologue.com>

Zarubin I. I. (1960) Shughni dictionary and texts. Moscow-Leningrad.