

Ethnologue: Languages of Tajikistan

Twenty-sixth edition data

David M. Eberhard, Gary F. Simons, and Charles D. Fennig, Editors

Based on information from the *Ethnologue*, 26th edition:

Eberhard, David M., Gary F. Simons, and Charles D. Fennig (eds.). 2023. *Ethnologue: Languages of the World*. Twenty-sixth edition. Dallas, Texas: SIL International. Online: <http://www.ethnologue.com>.

For personal use only

Permission to distribute or reuse this work (in whole or in part) may be obtained through the Copyright Clearance Center at <http://www.copyright.com>.

Contents

List of Abbreviations	3
How to Use This Digest	4
Country Overview	6
Language Status Profile	7
Statistical Summaries	8
Alphabetical Listing of Languages	10
Language Map	18
Languages by Population	19
Languages by Status	22
Languages by Region	26
Languages by Family	27
Language Code Index	28
Language Name Index	29
Bibliography	31

Copyright © 2023 by SIL International

All rights reserved. No part of this publication may be reproduced, redistributed, or transmitted in any form or by any means—electronic, mechanical, photocopying, recording, or otherwise—without the prior written permission of SIL International, with the exception of brief excerpts in articles or reviews.

List of Abbreviations

A	Agent in constituent word order
<i>alt.</i>	alternate name for
<i>alt. dial.</i>	alternate dialect name for
AOV	Agent-Object-Verb
C	Consonant in canonical syllable patterns
CDE	Convention against Discrimination in Education (1960)
<i>Class</i>	Language classification
CPPDCE	Convention on the Protection and Promotion of the Diversity of Cultural Expressions (2005)
CSICH	Convention for the Safeguarding of Intangible Cultural Heritage (2003)
DBS	Deaf Bible Society
<i>dial.</i>	primary dialect name for
DOOR	Deaf Opportunity Outreach
ICCPR	International Covenant on Civil and Political Rights (1966)
km	kilometer(s)
L1 / L2	first language / second (or other additional) language
<i>Lg Dev</i>	Language development
<i>Lg Use</i>	Language use
m	meter(s)
P	Patient in constituent word order
PARADISEC	Pacific And Regional Archive for Digital Sources In Endangered Cultures
<i>pej.</i>	pejorative
pl.	plural
S	Subject in constituent word order
sg.	singular
SIL	SIL International
SOV	Subject-Object-Verb
<i>Type</i>	Typological information
UNSD	United Nations Statistics Division
V	Vowel in canonical syllable patterns

How to Use This Digest

This *Ethnologue* country digest provides an extract of the information about the language situation in Tajikistan that is published in the 26th edition of *Ethnologue: Languages of the World* (see <http://www.ethnologue.com>), including some ways of presenting the information that are not available in the online version. The digest begins with a “Country Overview” (page 6) and “Statistical Summaries” (page 8) of languages and number of speakers by language size, by language status, and by language family.

The “Alphabetical Listing of Languages” (page 10) provides detailed information on the 28 languages listed in the *Ethnologue* for the country of Tajikistan. This includes languages that are indigenous to the country, languages that have developed well-established multigenerational speaker communities after immigrating in the past, and languages that have a significant presence in the country but are not established (that is, not being transmitted to the next generation within the country). A complete language entry has the following form and content:

Primary language name [ISO 639-3 code] (Alternate names). Autonym. *Users*: Country user population. Population stability comment. Population remarks. Monolingual population. Ethnic population. *Location*: Location. *Status*: EGIDS level. Special cases. Language function in country. *Class*: Linguistic classification. Macrolanguage membership. *Dialects*: Dialect names. Intelligibility and dialect relations. Lexical similarity. *Type*: Linguistic typology information. *Lg Use*: Remarks on use of the language. Domains of use. User age range. Language attitudes. Bilingualism remarks. Use as second language. *Lg Dev*: Literacy rates. Literacy remarks. Use in education. Publications and use in media. Revitalization efforts. Language development agencies. *DLS*: Digital support. *Writing*: Scripts used. *Other*: Non-indigenous. General remarks. Religion. *Worldwide*: Total population in all countries. Other countries where used.

See <http://www.ethnologue.com/methodology/#languagePages> for a full description of these information elements. If the autonym contains the “?” character, this indicates a complex non-Roman character that the PDF-creating software we are using is not able to render. We regret the inconvenience.

Many ways of finding languages are provided. “Languages by Population” (page 19) lists the languages in order of their first-language speaker populations. “Languages by Status” (page 22) lists the languages by their level of development or endangerment as measured on EGIDS, the Expanded Graded Intergenerational Disruption Scale (Lewis and Simons 2010). “Languages by Region” (page 26) lists the top-level administrative subdivisions of Tajikistan and the languages located within each. “Languages by Family” (page 27) lists the languages by their linguistic classifications. “Language Code Index” (page 28) gives an alphabetical listing of all the three-letter codes from ISO 639-3 that are used in this digest to uniquely identify languages. “Language Name Index” (page 29) lists every name that appears in the language listings as a primary or alternate name of a language or dialect. A total of 73 unique names are associated with the 28

languages described in this digest.

Finally, a listing of all the published sources cited within this digest is found in “Bibliography” (page 31). The published sources are cited using standard in-text citations enclosed in parentheses, consisting of the author’s or editor’s surname followed by the year of publication. Unpublished sources including personal communications and unpublished reports are also acknowledged when specific statements or facts are attributed to them. They are identified using in-text citations enclosed in parentheses in which the year of the communication is given first, followed by the source’s first initial and surname. In such a case, there is no corresponding entry in the bibliography.

This digest is designed for use in both digital and print formats. The cross-references are thus rendered as page numbers that are hyperlinks. When using the document in printed form, simply turn to the referenced page by number. When using it in digital form, click on the blue text to jump to the cross-referenced location.

If you believe any of the information about a language in this digest is in error or if you are able to supply missing information, please send your proposed change to the editor using one of the means given below. Provide as much information as possible about the source of your information. Full bibliographic details of published sources are especially helpful.

The preferred method of submitting corrections and additions is to join our Contributor Program by applying at <http://www.ethnologue.com/contributor-program>. With a contributor account you will be entitled to complimentary access to the website and will be able to use the Contribute tab on the page for a language or country in order to propose corrections and additions. The advantage of giving feedback in this way is that it becomes part of the public record on the website. You will also be automatically notified of the editorial action.

Alternatively, you may submit corrections and additions by e-mail to:

Ethnologue_Editor@sil.org

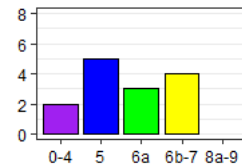
Or by post to:

Editor, Ethnologue
SIL International
7500 West Camp Wisdom Road
Dallas, TX 75236, U.S.A.

Country Overview

Name of country	Tajikistan
Other names	Republic of Tajikistan
Population	9,506,000 (2020 census)
Principal languages	Russian, Tajik
Literacy rate	100% (2015 World Factbook)
International conventions	CDE (1992), CPPDCE (2007), CSICH (2010), ICCPR (1999)
General references	Akiner 1983, Comrie 1987, Kibrik 1991, Sebeok 1963

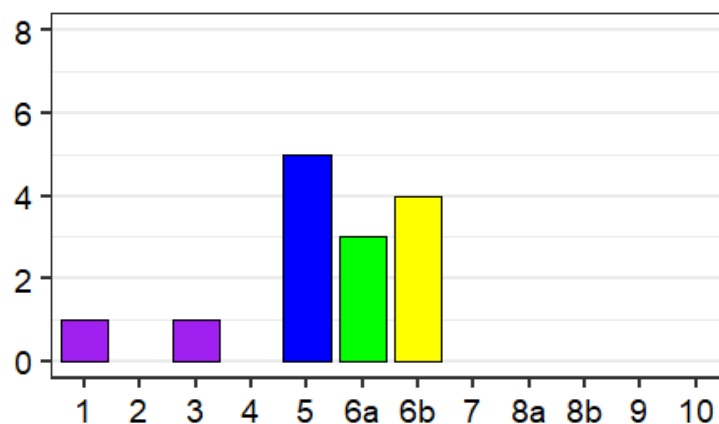
Language counts The number of established languages listed for Tajikistan is 14. All are living languages. Of these, 11 are indigenous and 3 are non-indigenous. Furthermore, 2 are institutional, 5 are developing, 3 are vigorous, and 4 are in trouble. Also listed are 14 unestablished languages.



See the next page for an explanation of the summary categories for language vitality used in the above counts and graph.

Language Status Profile

The following histogram gives a graphic profile of the established languages in Tajikistan with respect to their status of language development versus language endangerment. This includes all of the languages appearing in the “Alphabetical Listing of Languages” (page 10) that report an EGIDS level after *Status*; macrolanguages and unestablished languages are not included in the profile. The horizontal axis plots the estimated level of development or endangerment as measured on the EGIDS scale (Lewis and Simons 2010). The height of each bar indicates the number of languages that are estimated to be at the given level. Consult “Languages by Status” (page 22) in order to see the specific languages for each level.



The color coding of the bars in the histogram above matches the color scheme used in the summary profile graph on the preceding page. In this scheme, the EGIDS levels are grouped as follows:

- Purple = Institutional (EGIDS 0–4) — The language has been developed to the point that it is used and sustained by institutions beyond the home and community.
- Blue = Developing (EGIDS 5) — The language is in vigorous use, with literature in a standardized form being used by some though this is not yet widespread or sustainable.
- Green = Vigorous (EGIDS 6a) — The language is in vigorous use among all generations and remains unstandardized.
- Yellow = In trouble (EGIDS 6b–7) — Intergenerational transmission is in the process of being broken, but the child-bearing generation can still use the language so it is possible that revitalization efforts could restore transmission of the language in the home.
- Red = Dying (EGIDS 8a–9) — The only fluent users (if any) are older than child-bearing age, so it is too late to restore natural intergenerational transmission through the home; a mechanism outside the home would need to be developed.
- Black = Extinct (EGIDS 10) — The language is no longer used and no one retains a sense of ethnic identity associated with the language.

Statistical Summaries

The “Alphabetical Listing of Languages” (page 10) provides a detailed listing of all the languages of Tajikistan. This section steps back from the detail to offer a summary view of the language situation in the country. Specifically, it offers three numerical tabulations of the living established languages of Tajikistan and their users: by language size, by language status, and by language family.

Summary by language size

Table 1 summarizes the distribution of living established languages in Tajikistan by number of L1 speakers. The *Population range* column categorizes the sizes of the languages by order of magnitude (in terms of the number of digits in the population of first-language speakers). Consult “Languages by Population” (page 19) for a listing of the specific languages in each range category.

The *Count* column gives the number of living established languages within the specified population range. The *Percent* column gives the share of the count for that population range as a percentage of the total number of languages given at the bottom of the Count column. The *Cumulative* column gives the cumulative sum of the percentage of languages going from top to bottom in the column.

The *Total* column gives the total L1 population of all the languages in the given range category. The second *Percent* column gives the percentage of the total country population as estimated at the bottom of the Total column. Note that if the table has a row for Unknown, representing languages for which the *Ethnologue* does not have a population estimate, the calculation of population percentage is not able to take those languages into account. The final *Cumulative* column gives the cumulative sums of the population percentages going from top to bottom in the column.

Table 1: Distribution of languages by number of first-language speakers

Population range	Living languages			Number of speakers		
	Count	Percent	Cumulative	Total	Percent	Cumulative
1,000,000 to 9,999,999	2	14.3	14.3%	7,440,000	95.33573	95.33573%
10,000 to 99,999	8	57.1	71.4%	351,600	4.50538	99.84111%
1,000 to 9,999	3	21.4	92.9%	11,900	0.15249	99.99359%
100 to 999	1	7.1	100.0%	500	0.00641	100.00000%
<i>Totals</i>	14	100.0		7,804,000	100.00000	

Summary by language status

Table 2 summarizes the distribution of living established languages in Tajikistan by their status in terms of language development or language endangerment. The *EGIDS* column categorizes the

languages by their level on the EGIDS scale. Consult “Languages by Status” (page 22) for a listing of the specific languages that have been assigned to each level. Note that the EGIDS level reported here is for the status of the language in Tajikistan. Languages that are also used in other countries may be assigned to a different EGIDS level in those countries.

The next six columns are as in Table 1. In addition, the *Mean* column gives the average L1 population of all the languages with the given EGIDS level and the *Median* column gives the median L1 population for the languages at that level, that is, half of the languages at that level have a higher population and half have a lower population. If there are any languages with an unknown population, these are ignored in the calculation of the mean and the median.

Table 2: Distribution of languages by vitality status

EGIDS	Living languages			Number of speakers				
	<i>Count</i>	<i>Percent</i>	<i>Cumulative</i>	<i>Total</i>	<i>Percent</i>	<i>Cumulative</i>	<i>Mean</i>	<i>Median</i>
1	1	7.1	7.1%	6,380,000	81.7529	81.7529%	6,380,000	6,380,000
3	1	7.1	14.3%	44,600	0.5715	82.3244%	44,600	44,600
5	5	35.7	50.0%	1,300,000	16.6581	98.9826%	260,000	71,400
6a	3	21.4	71.4%	42,100	0.5395	99.5220%	14,033	1,600
6b	4	28.6	100.0%	37,300	0.4780	100.0000%	9,325	10,500
<i>Totals</i>	14	100.0		7,804,000	100.0000			

Summary by language family

The genealogical classifications given in the language entries of the “Alphabetical Listing of Languages” (page 10) name 4 different top-level groups. Table 3 summarizes the distribution of living established languages and their L1 populations within these families. The columns are as for table 2, with the exception that *Cumulative* is excluded since there is no inherent ordering of the families.

Table 3: Distribution of languages by language family

Language family	Living languages		Number of speakers				
	<i>Count</i>	<i>Percent</i>	<i>Total</i>	<i>Percent</i>	<i>Mean</i>	<i>Median</i>	
Afro-Asiatic	1	7.1	1,300	0.0	1,300	1,300	
Indo-European	10	71.4	6,624,300	84.9	662,430	23,700	
Sign language	1	7.1	47,000	0.6	47,000	47,000	
Turkic	2	14.3	1,131,400	14.5	565,700	565,700	
<i>Totals</i>	14	100.0	7,804,000	100.0			

Alphabetical Listing of Languages

Aimaq [aiq]. *Users*: 7,900 in Tajikistan (2017). *Location*: Khatlon region: west of Vakhsh river, Jamshidi and Khazara area. *Status*: Unestablished. *Class*: Indo-European, Indo-Iranian, Iranian, Western, Southwestern, Persian. *Writing*: Arabic script, Naskh variant. *Other*: Non-indigenous. *Worldwide*: Total users in all countries: 1,891,900. Global EGIDS level: 5 (Developing). Indigenous in: Afghanistan. Also established in: Iran, Turkmenistan.

Arabic, Tajiki Spoken [abh] (Arabi, Bukhara Arabic, Buxara Arabic, Central Asian, Jugari, Tajiki Arabic). *Users*: 1,300 in Tajikistan (2017). *Location*: Khatlon province: Vakhsh valley, mainly small villages. *Status*: 6b* (Threatened). *Class*: Afro-Asiatic, Semitic, Central, South, Arabic. A member of macrolanguage Arabic [ara] (Saudi Arabia). *Lg Use*: Only a few speak Arabic. Shifting to Tajiki [tgk]. No diglossia with Modern Standard Arabic [arb]. Also use Tajik [tgk]. *DLS*: Emerging (0.01). *Writing*: Arabic script, Naskh variant. *Other*: Endogamous, do not mix with speakers of other languages. Muslim. *Map*: 18. *Worldwide*: Total users in all countries: 17,300. Also indigenous in: Afghanistan.

Armenian, Western [hyw]. *Users*: 8,900 in Tajikistan (2017). *Location*: Dushanbe province: Dushanbe; Sughd province: Chkalovsk and Khujand; Khatlon province: Qurghorteppa. Small communities in urban areas. *Status*: Unestablished. *Class*: Indo-European, Armenian. *Lg Dev*: Fully developed. *Writing*: Armenian script. *Other*: Non-indigenous. *Worldwide*: Total users in all countries: 1,582,290 (as L1: 1,576,490; as L2: 5,800). Global EGIDS level: 5 (Developing). Indigenous in: Armenia, Georgia, Turkey. Also established in 21 other countries and unestablished in 13 more.

Azerbaijani, North [azj]. Autonym: Azərbaycan dili, Azərbaycanca. *Users*: 5,640 in Tajikistan (2017). *Status*: Unestablished. *Class*: Turkic, Southern, Azerbaijani. *Type*: SOV; postpositions; no articles; case-marking (6 cases); verb affixes mark person, number; passives; tense; 24 consonant and 9 vowel phonemes; non-tonal; stress on final syllable; vowel harmony. *Lg Dev*: Literature. Radio. TV. Dictionary. Grammar. Bible: 1891–2009. *Writing*: Arabic script, Naskh variant, used until 1920s. Braille script. Cyrillic script, official usage in Dagestan. Latin script, official usage in Azerbaijan since 1992. *Other*: Non-indigenous. *Worldwide*: Total users in all countries: 9,110,020. Global EGIDS level: 1 (National). Indigenous in: Armenia, Azerbaijan, Georgia. Also established in: Russian Federation. Unestablished in: Afghanistan, Estonia, Germany, Kazakhstan, Kyrgyzstan, Latvia, Sweden, Turkmenistan, Ukraine, Uzbekistan.

Balochi, Western [bgn]. *Users*: 6,350 in Tajikistan (2017). *Location*: Scattered. *Status*: Unestablished. *Class*: Indo-European, Indo-Iranian, Iranian, Western, Northwestern, Balochi. *Lg Dev*: Newspapers. Radio. Dictionary. Grammar. Texts. Bible portions: 1984–2005. *Writing*: Arabic script, Naskh variant. Arabic script, Nastaliq variant, primary usage. Cyrillic script, used in Turkmenistan. *Other*: Non-indigenous. *Worldwide*: Total users in all countries: 2,283,350. Global EGIDS level: 5 (Developing). Indigenous in: Afghanistan, Iran, Pakistan. Also established in: Turkmenistan.

Bashkort [bak]. *Users:* 8,200 in Tajikistan (2003 J. Leclerc). *Location:* Scattered. *Status:* Unestablished. *Class:* Turkic, Western, Uralian. *Type:* SOV; postpositions; noun head final; no articles; case-marking (6 cases); verb affixes mark person, number; 27 consonants, 9 vowels; non-tonal; stress on final syllable; vowel harmony. *Lg Dev:* Literature. Dictionary. Grammar. NT: 2015. *Writing:* Arabic script, Naskh variant, no longer in use. Cyrillic script. *Other:* Non-indigenous. *Worldwide:* Total users in all countries: 1,180,940. Global EGIDS level: 2 (Provincial). Indigenous in: Russian Federation. Unestablished in: Kazakhstan, Kyrgyzstan, Turkmenistan, Ukraine, Uzbekistan.

Belarusian [bel]. *Users:* 6,430 in Tajikistan (2017). *Status:* Unestablished. *Class:* Indo-European, Balto-Slavic, Slavic, East. *Type:* SVO; prepositions; noun head final; gender (masculine/feminine/neuter); no articles; case-marking (6 cases); verb affixes mark person, number, gender of subject (in past tense); passives; tense and aspect; 37 consonant and 6 vowel phonemes; non-tonal; free stress. *Lg Dev:* Dictionary. Grammar. Bible: 1973–2017. *Writing:* Arabic script, Naskh variant, developed in the 16th century. Braille script, no longer in use. Cyrillic script, official usage. Latin script, official usage 1941–1944, modern occasional usage. *Other:* Non-indigenous. *Worldwide:* Total users in all countries: 3,766,550 (as L1: 1,376,550; as L2: 2,390,000). Global EGIDS level: 5 (Developing). Indigenous in: Belarus. Also established in: Lithuania, Poland, Ukraine. Unestablished in: Argentina, Azerbaijan, Czechia, Estonia, Germany, Kazakhstan, Kyrgyzstan, Latvia, Russian Federation, Spain, Sweden, Turkmenistan, United States, Uzbekistan.

Chuvash [chv]. *Users:* 3,940 in Tajikistan (2017). *Status:* Unestablished. *Class:* Turkic, Bolgar. *Type:* SOV. *Lg Dev:* Literature. Newspapers. Radio. TV. Dictionary. Grammar. Bible: 2010. *Writing:* Cyrillic script. *Other:* Non-indigenous. *Worldwide:* Total users in all countries: 738,150. Global EGIDS level: 6b (Threatened). Indigenous in: Russian Federation. Unestablished in: Kazakhstan, Latvia, Turkmenistan, Ukraine, Uzbekistan.

Erzya [myv]. *Users:* 8,800 in Tajikistan (2017). *Status:* Unestablished. *Class:* Uralic, Mordvin. *Lg Dev:* Dictionary. Grammar. NT: 1827–2006. *Writing:* Cyrillic script. *Other:* Non-indigenous. *Worldwide:* Total users in all countries: 245,780. Global EGIDS level: 5 (Developing). Indigenous in: Russian Federation. Also established in: Armenia. Unestablished in: Azerbaijan, Belarus, Kazakhstan, Kyrgyzstan, Latvia, Turkmenistan, Ukraine, Uzbekistan.

Georgian [kat]. *Users:* 1,400 in Tajikistan (2017). *Location:* Scattered. *Status:* Unestablished. *Class:* Kartvelian, Georgian. *Type:* SOV; postpositions; noun head final; no articles; case-marking (7 cases); verb affixes mark person, number, object; ergativity; passives; tense and aspect; 27 consonants, 5 vowels, no diphthongs; non-tonal; stress on first syllable; long consonant clusters. *Lg Dev:* Fully developed. Bible: 1743–2021. *Writing:* Braille script. Georgian (Mkhedruli and Mtavruli) script, primary usage. Khutsuri (Asomtavruli and Nuskhuri) script, no longer in use. *Other:* Non-indigenous. *Worldwide:* Total users in all countries: 3,878,780 (as L1: 3,724,240; as L2: 154,540). Global EGIDS level: 1 (National). Indigenous in: Georgia, Turkey. Also established in: Azerbaijan, Iran. Unestablished in: Armenia, Canada, Germany, Greece, Kazakhstan, Kyrgyzstan, Latvia, Russian Federation,

Spain, Sweden, Turkmenistan, Ukraine, Uzbekistan.

Hazaragi [haz]. *Users*: 1,700 in Tajikistan (2022 Joshua Project), based on ethnicity. *Status*: Unestablished. *Class*: Indo-European, Indo-Iranian, Iranian, Western, Southwestern, Persian. *Lg Dev*: Radio. Dictionary. Bible portions: 2008–2015. *Writing*: Arabic script, Naskh variant. Arabic script, Nastaliq variant, used in handwriting and for book titles. *Other*: Non-indigenous. *Worldwide*: Total users in all countries: 3,660,000. Global EGIDS level: 5 (Developing). Indigenous in: Afghanistan. Also established in: Iran, Pakistan. Unestablished in: Australia.

Ishkashimi [isk] (Ishkashim). *Users*: 500 in Tajikistan (1990 A. Kibrik). *Location*: Kuhistoni Badakhshon region: Ryn and Sumjin villages, near Ishkashim. *Status*: 6a (Vigorous). *Class*: Indo-European, Indo-Iranian, Iranian, Eastern, Southeastern, Pamir. *Lg Use*: Ishkashimi is very vital. This is because the fathers very much insist on their children acquiring Ishkashimi as their first language when they grow up. Home, village. Used by all. Positive attitudes. *Lg Dev*: Grammar. *Writing*: Unwritten. *Other*: Muslim. *Map*: 18. *Worldwide*: Total users in all countries: 3,000. Also indigenous in: Afghanistan.

Kyrgyz [kir]. *Users*: 71,400 in Tajikistan (2017). Ethnic population: 152,000 (2017). *Location*: Kuhistoni Badakhshon region and Regions of Republican Subordination. *Status*: 5* (Dispersed). *Class*: Turkic, Western, Aralo-Caspian. *Type*: SOV; postpositions; no articles; case-marking (7 cases); verb affixes mark person, number; passives; tense; 19 consonant and 8 vowel phonemes; non-tonal; stress on final syllable; vowel harmony. *Lg Dev*: Newspapers. Radio. TV. Videos. Dictionary. Grammar. Bible: 1995–2004. *Writing*: Arabic script, Naskh variant, used in China. Cyrillic script, not used in Afghanistan. Latin script, used in Turkey. *Map*: 18. *Worldwide*: Total users in all countries: 5,154,500. Global EGIDS level: 1 (National). Also indigenous in: China, Kyrgyzstan. Also established in: Afghanistan, Turkey. Unestablished in: Germany, Kazakhstan, Russian Federation, Turkmenistan, Uzbekistan.

Lak [lbe]. *Users*: 2,100 in Tajikistan (2017). *Status*: Unestablished. *Class*: Nakh-Daghestanian, Lak. *Lg Dev*: Literature. Newspapers. Periodicals. Radio. Videos. Dictionary. Grammar. Bible portions: 1996–2007. *Writing*: Arabic script, Naskh variant, used until 1928. Cyrillic script. Latin script, used between 1928-1938. *Other*: Non-indigenous. *Worldwide*: Total users in all countries: 100,000. Global EGIDS level: 4 (Educational). Indigenous in: Russian Federation. Unestablished in: Azerbaijan, Turkmenistan, Ukraine, Uzbekistan.

Parya [paq] (Afghana-Yi Nasfurush, Afghana-Yi Siyarui, Changgars, Laghmani, Pbharya). *Users*: 1,600 in Tajikistan (2017). *Location*: Regions of Republican Subordination: Hissar valley. *Status*: 6a* (Vigorous). *Class*: Indo-European, Indo-Iranian, Indo-Aryan, Intermediate Divisions, Western, Unclassified. *Dialects*: May be Marwari [rwr] dialect, related to Punjabi [pan], or Laghman dialect of Southeast Pashai [psi] of Afghanistan. Subgroups: Kalu, Jitain, Juni, Maggar, Bisiyan, Mussali, Shuiya. *Lg Use*: Despite small population, it is spoken more vigorously amongst themselves than larger language groups. Uzbek and Tajiki men who marry Parya women learn Parya and become assimilated into the community. Home. Also use Tajik [tgk]. *DLS*: Still. *Writing*: Unwritten. *Other*: Came in 1880 to Tajikistan from Laghman,

Afghanistan. Subgroup names similar to Changgars of Lahore, and to names used for groups mistakenly called Roma. Muslim. *Map*: 18. *Worldwide*: Total users in all countries: 2,600. Also established in: Afghanistan. Unestablished in: Uzbekistan.

Pashto, Southern [pbt]. *Users*: 32,400 in Tajikistan (2017). *Location*: Khatlon region. *Status*: 5* (Dispersed). *Class*: Indo-European, Indo-Iranian, Iranian, Eastern, Southeastern, Pashto. *Type*: SOV; noun head final; gender (masculine/feminine); no articles; case-marking (3 cases); verb affixes mark person, number, gender of subject; passives; tense and aspect; causatives. *Lg Dev*: Literature. Newspapers. Periodicals. Radio. TV. Videos. Dictionary. Grammar. Texts. Bible portions: 2011–2019. *Writing*: Arabic script, Naskh variant, used in print. Arabic script, Nastaliq variant, used in handwriting and for book titles. *Map*: 18. *Worldwide*: Total users in all countries: 19,871,430 (as L1: 16,391,430; as L2: 3,480,000). Global EGIDS level: 1 (National). Also indigenous in: Afghanistan, Iran, Pakistan. Unestablished in: Belgium, Finland, Germany, Sweden, Turkey, Turkmenistan, United Arab Emirates, United Kingdom.

Persian, Iranian [pes] (Persian, Western Farsi). *Users*: 89,200 in Tajikistan (2017). *Location*: Khatlon region: various economic development sites; Dushanbe. *Status*: 5* (Dispersed). *Class*: Indo-European, Indo-Iranian, Iranian, Western, Southwestern, Persian. *Type*: SOV; noun head initial; no articles; tense and aspect; causatives; 23 consonant and 6 vowel phonemes; stress on final syllable (or root). *Lg Dev*: Fully developed. Bible: 1845–1995. *Writing*: Arabic script, Naskh variant. Arabic script, Nastaliq variant, used in Iran. *Other*: Non-indigenous. Not a literary language in Tajikistan. Muslim. *Worldwide*: Total users in all countries: 78,623,350 (as L1: 57,192,350; as L2: 21,431,000). Global EGIDS level: 1 (National). Indigenous in: Iran. Also established in: Bahrain, Iraq, Oman, Qatar, United Arab Emirates. Unestablished in: Australia, Austria, Azerbaijan, Belgium, Canada, Denmark, Finland, France, Germany, India, Japan, Kuwait, Netherlands, New Zealand, Norway, Pakistan, Russian Federation, Saudi Arabia, Sweden, Syria, Turkey, Turkmenistan, United Kingdom, United States, Uzbekistan, Yemen.

Russian [rus] (Rusija). *Users*: 44,600 in Tajikistan (2017). Ethnic population: 303,000 (2017). *Location*: Widespread. *Status*: 3 (Wider communication). De facto national working language, no longer language of record in higher levels of government. *Class*: Indo-European, Balto-Slavic, Slavic, East. *Type*: SVO; prepositions; genitives after noun heads; adjectives, numerals before noun heads; question word initial; 1 prefix on a word; recursive addition of suffixes allowed; gender (masculine/feminine/neuter); no articles; case-marking (6 cases); verb affixes mark person, number; passives; tense and aspect; comparatives; 32 consonants, 5 vowels, 4 diphthongs; non-tonal; free stress. *Lg Use*: Used as L2 by Russian Sign Language [rsl], Tajik [tgk], Yagnobi [yai]. *Lg Dev*: Taught as subject in primary and secondary schools from grade 2. Fully developed. Bible: 1876–2011. *Writing*: Braille script. Cyrillic script, primary usage. *Other*: Non-indigenous. *Worldwide*: Total users in all countries: 254,997,130 (as L1: 146,954,150; as L2: 107,987,980). Global EGIDS level: 0 (International). Indigenous in: Belarus, Georgia, Kazakhstan, Russian Federation, Ukraine. Also established in 16 other countries and unestablished in 30 more.

Russian Sign Language [rsl] (RSL, TSL, Tajik Sign Language). *Users*: 47,000 in Tajikistan (2021 DBS/DOOR/SIL). Estimated 28,000–56,000 signing deaf, assuming 0.3%–0.6% of total population. *Location*: Scattered, with concentrations around Dushanbe. *Status*: 5 (Developing). *Class*: Sign language, Deaf community sign language. *Dialects*: None known. Includes some local signs, and some grammatical differences from standard RSL that suggest it may need to be considered a separate language. Considerable variation within Tajikistan, since many deaf have never attended school and have had little exposure to the standard variety (2019 J. Power). *Type*: One-handed fingerspelling (Cyrillic script). *Lg Use*: Vigorous. Used by all. Some also use Russian [rus]. Deaf people born before about 2000 learned Russian in school. Some also use Tajik [tgk]. Deaf people born after about 2000 have been taught Tajik in schools, and the government currently promotes Tajik over Russian. *Lg Dev*: TV. Videos. Dictionary. *Other*: Non-indigenous. Two schools for the deaf in Dushanbe area (grades 1–10), another in Khujand, with a mixture of oralist approaches and some use of signs (2019 J. Power). Muslim. *Worldwide*: Total users in all countries: 909,000. Indigenous in: Russian Federation. Also established in: Armenia, Belarus, Estonia, Kazakhstan, Kyrgyzstan, Moldova, Turkmenistan, Uzbekistan. Unestablished in: Israel.

Shughni [sgh] (Shugnan-Rushan). Autonym: Khugnone. *Users*: 40,000 in Tajikistan (1975 SIL). 20,000 Shugan, 1,500–2,000 Oroshor, 15,000 Rushan. Ethnic population: 73,800 (1990 A. Kibrik). Includes 50,000 Shughni, 2,000 Oroshor, 18,000 Rushan, 800 Khufi, 3,000 Bartang. *Location*: Kuhistoni Badakhshon region: Pamir mountains. *Status*: 6a (Vigorous). *Class*: Indo-European, Indo-Iranian, Iranian, Eastern, Southeastern, Pamir, Shughni-Yazgulami. *Dialects*: Rushani (Oroshani, Roshani, Rushan), Bartangi (Bartang), Oroshor (Roshorvi), Khufi (Chuf, Khuf), Shughni (Khugni, Shighni, Shugan, Shugnan). Khufi and Bartangi dialects may be separate languages. Oroshani may be separate from Rushani. Not intelligible of Sarikoli [srh] (called, Tajiki, in China). *Lg Use*: Home. Used by all. Positive attitudes. Also use Tajik [tgk], as a literary language. Used as L2 by Wakhi [wbl]. *Lg Dev*: Literature. Dictionary. Texts. Bible portions: 2015. *DLS*: Emerging (0.06). *Writing*: Arabic script, Naskh variant, used in Afghanistan. Cyrillic script, used in Tajikistan. Latin script, used in Tajikistan. *Other*: Traditional territory and way of life. Muslim. *Map*: 18. *Worldwide*: Total users in all countries: 80,000. Global EGIDS level: 5 (Developing). Also indigenous in: Afghanistan.

Tajik [tgk] (Tadzhik, Tajiki, Tajiki Persian, “Galcha” *pej*). *Users*: 6,380,000 in Tajikistan (2012 UNSD). Ethnic population: 7,520,000 (2017). *Location*: Widespread. *Status*: 1 (National). Statutory national language (1994, Constitution, Article 2). *Class*: Indo-European, Indo-Iranian, Iranian, Western, Southwestern, Persian. *Dialects*: 4 groups of small dialects; no distinct boundaries. Dialect blending into Dari Persian [prs] in Afghanistan. *Type*: SOV; verb affixes mark person, number; tense and aspect; causatives; 27 consonant and 6 vowel phonemes; nono-tonal; stress on final syllable. *Lg Use*: Also use Northern Uzbek [uzn]. Also use Russian [rus]. *Lg Dev*: Taught in most primary and secondary schools. Literature. Radio. Bible: 1992–1999. *DLS*: Vital (0.68). *Writing*: Arabic script, Naskh variant, used until 1928. Cyrillic script, used since 1940, primary usage. Hebrew script, used by Bukharan Jews in Uzbekistan. Latin script, used between 1928–1940. *Other*: Russian sources refer to Persian

dialects in Afghanistan as Tajiki. So-called Tajiki in China is actually Shugni [sgh] or Wakhi [wbl]. Some Tajiki-speaking Roma communities in Russian central Asia. Muslim. *Map*: 18. *Worldwide*: Total users in all countries: 8,287,420. Also indigenous in: Uzbekistan. Also established in: China, Kazakhstan, Kyrgyzstan. Unestablished in: Germany, Iran, Russian Federation, Turkmenistan, Ukraine.

Tatar [tat]. *Users*: 4,200 in Tajikistan (2012 UNSD). Ethnic population: 107,000 (2017). *Location*: Kuhistoni Badakhshon region: scattered. *Status*: Unestablished. *Class*: Turkic, Western, Uralian. *Type*: SOV; case-marking (6 cases); verb affixes mark person, number; passives; tense; causatives; 21 consonant and 12 vowel phonemes; non-tonal; stress on final syllable; vowel harmony; evidentiality. *Lg Dev*: Literature. Newspapers. Periodicals. Radio. TV. Dictionary. Grammar. Bible: 2016. *Writing*: Arabic script, Naskh variant, used prior to 1928. Cyrillic script. Latin script. *Other*: Non-indigenous. *Worldwide*: Total users in all countries: 5,127,970 (as L1: 4,317,970; as L2: 810,000). Global EGIDS level: 2 (Provincial). Indigenous in: Russian Federation. Also established in: China, Kazakhstan, Turkey. Unestablished in: Azerbaijan, Belarus, Estonia, Georgia, Kyrgyzstan, Latvia, Lithuania, Romania, Turkmenistan, Ukraine, United States, Uzbekistan.

Turkish [tur]. Autonym: Türkçe. *Users*: 1,000 in Tajikistan (2017). *Status*: Unestablished. *Class*: Turkic, Southern, Turkish. *Type*: SOV; postpositions; noun head final; indefinite articles; case-marking (6 cases); verb affixes mark person, number; passives; tense; causatives; comparatives; 19 consonant and 8 vowel phonemes; non-tonal; stress on final syllable; vowel harmony; evidentiality. *Lg Dev*: Fully developed. Bible: 1827–2006. *Writing*: Arabic script, Naskh variant, no longer in use. Braille script. Cyrillic script, used in Bulgaria. Greek script, no longer in use. Latin script, used since 1928 or 1929, primary usage. *Other*: Non-indigenous. *Worldwide*: Total users in all countries: 90,028,000 (as L1: 84,010,500; as L2: 6,017,500). Global EGIDS level: 1 (National). Indigenous in: Bulgaria, Cyprus, Greece, Turkey. Also established in: Germany, Jordan, Kazakhstan, Lebanon, North Macedonia, Romania, Serbia, Syria, Uzbekistan. Unestablished in: Albania, Australia, Austria, Azerbaijan, Belgium, Bosnia and Herzegovina, Brazil, Canada, Croatia, Denmark, Egypt, Finland, France, Georgia, Iran, Iraq, Israel, Italy, Japan, Kuwait, Kyrgyzstan, Libya, Morocco, Netherlands, New Zealand, Norway, Pakistan, Poland, Russian Federation, Saudi Arabia, Spain, Sweden, Switzerland, Turkmenistan, Ukraine, United Arab Emirates, United Kingdom, United States.

Turkmen [tuk]. *Users*: 13,700 in Tajikistan (2012 UNSD). Ethnic population: 32,500 (2017). *Location*: Khatlon region: scattered. *Status*: Unestablished. *Class*: Turkic, Southern, Turkmenian. *Type*: SOV; postpositions; no articles; case-marking (5 cases); verb affixes mark person, number; passives; tense; 23 consonant and 9 vowel phonemes; non-tonal; vowel harmony; evidentials. *Lg Dev*: Literature. Periodicals. Radio. TV. Dictionary. Grammar. Texts. Bible: 2017. *Writing*: Arabic script, Naskh variant, used in Afghanistan and Iran. Cyrillic script, used in Turkmenistan, not used in Afghanistan. Latin script, official usage in Turkmenistan since 1991. *Other*: Non-indigenous. *Worldwide*: Total users in all countries: 6,562,260. Global EGIDS level: 1 (National). Indigenous in: Afghanistan, Iran, Turkmenistan. Also established in: Iraq, Russian Federation, Turkey. Unestablished in: Germany, Kazakhstan,

Kyrgyzstan, Pakistan, Ukraine, Uzbekistan.

Uzbek, Northern [uzn] (o‘zbek, o‘zbek tili, o‘zbekcha). *Users*: 1,060,000 in Tajikistan (2017). Ethnic population: 1,230,000 (2017). *Location*: Khatlon and Sughd regions; Regions of Republican Subordination. *Status*: 5* (Dispersed). *Class*: Turkic, Eastern. *Type*: SOV; postpositions; noun head final; no articles; case-marking (5 cases); verb affixes mark person, number; passives; tense; comparatives; 24 consonant and 8 vowel phonemes; non-tonal; stress on final syllable; has lost its historical vowel harmony and its vowel system now resembles that of Tajiki. *Lg Use*: Used as L2 by Tajik [tgk]. *Lg Dev*: Radio. TV. Dictionary. Grammar. Bible: 2016. *Writing*: Arabic script, Naskh variant, symbolic use in Uzbekistan, sometimes used in China. Braille script. Cyrillic script, official usage in Uzbekistan between 1940–1992, continued widespread use in Uzbekistan and China. Latin script, official usage in Uzbekistan since 1992. *Map*: 18. *Worldwide*: Total users in all countries: 27,937,610 (as L1: 27,936,530; as L2: 1,080). Global EGIDS level: 1 (National). Also indigenous in: Turkmenistan, Uzbekistan. Also established in: China, Kazakhstan, Kyrgyzstan. Unestablished in: Canada, Finland, Germany, Russian Federation, South Korea, Sweden, Turkey, Ukraine.

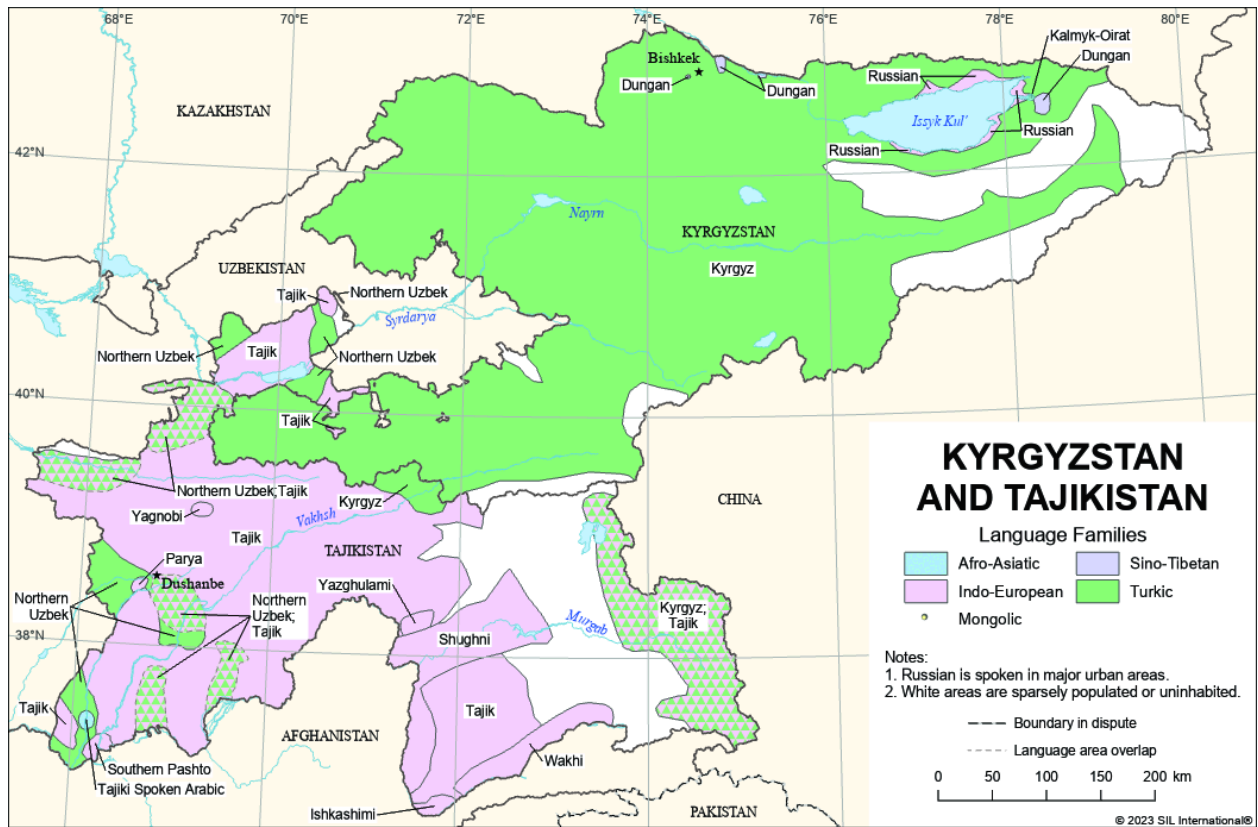
Wakhi [wbl] (Guhjali, Khik, Khikwar, Vakhan, Wakhani, Wakhigi). *Users*: 15,000 in Tajikistan (2012 M. Merritt). Ethnic population: 20,000 (1990 A. Kibrik). *Location*: Kuhistoni Badakhshon region: Pamir mountains. *Status*: 6b (Threatened). *Class*: Indo-European, Indo-Iranian, Iranian, Eastern, Southeastern, Pamir. *Dialects*: Western Wakhi, Central Wakhi, Eastern Wakhi. *Type*: SOV. *Lg Use*: Some domains. Some of all ages. Positive attitudes. Many also use Shughni [sgh]. Also use Tajik [tgk], as literary language. *Lg Dev*: Radio. Dictionary. Grammar. Texts. Bible portions: 2001. *Writing*: Arabic script, Naskh variant, used in Afghanistan. Arabic script, Nastaliq variant, used in handwriting and for book titles. Cyrillic script, used in Tajikistan. Latin script, used in Pakistan, preferred script, not firmly established. *Other*: Muslim. *Map*: 18. *Worldwide*: Total users in all countries: 52,200. Global EGIDS level: 5 (Developing). Also indigenous in: Afghanistan, Pakistan. Also established in: China.

Yagnobi [yai] (Soghdi, Yaghnabi, Yaghnobi, Yaghnubi, Yagnabi, Yagnob, Yagnubi). *Users*: 12,000 (2004 SIL). *Location*: Sughd region: Zafarabad; north of Dushanbe; Yagnob river (the homeland), high mountain valley. *Status*: 6b* (Threatened). *Class*: Indo-European, Indo-Iranian, Iranian, Eastern, Northeastern. *Dialects*: Western Yagnobi, Eastern Yagnobi. *Lg Use*: Also use Russian [rus], as a literary language. Also use Tajik [tgk], as a literary language. *Lg Dev*: Grammar. *DLS*: Emerging (0.01). *Writing*: Cyrillic script. *Other*: Were resettled in 1970s to Zafarabad in north Tajikistan on the border with Uzbekistan; changed from cool mountain valleys with water to irrigated desert plains. Muslim. *Map*: 18.

Yazghulami [yah] (Iazgulem, Yazghulomi, Yazgulam, Yazgulami, Yazgulyam, Yazgulyami, Yuzdomi zavég, Zgyamigyai). *Users*: 9,000 (2010). Ethnic population: 13,000. *Location*: Kuhistoni Badakhshon province: Andarbag, Budun, Khavavd, Motravn, Shavud, Ubaghn, Vishkharv, Zaich, Zhafag, and Zhamag in Yazgulyam river valley; Dushanbe and Kuybishev areas. *Status*: 6b (Threatened). *Class*: Indo-European, Indo-Iranian, Iranian, Eastern, Southeastern, Pamir, Shugni-Yazgulami. *Dialects*: Upper Yazghulami, Lower Yazghulami.

Little dialect difference. *Lg Use*: Home. Some of all ages. Positive attitudes. Also use Tajik [tgk]. *Lg Dev*: Dictionary. Texts. *DLS*: Still. *Writing*: Cyrillic script. Latin script. *Other*: Muslim. *Map*: 18.

Language Map



Languages by Population

In this section the languages of Tajikistan are listed in order of their population of first-language speakers within the country, from highest to lowest. The entries report just the population and status elements.

1,000,000 to 9,999,999

Tajik [tgk] *Users:* 6,380,000 in Tajikistan (2012 UNSD). *Ethnic population:* 7,520,000 (2017). *Status:* 1 (National). *Statutory national language* (1994, Constitution, Article 2). *Worldwide:* Total users in all countries: 8,287,420.

Uzbek, Northern [uzn] *Users:* 1,060,000 in Tajikistan (2017). *Ethnic population:* 1,230,000 (2017). *Status:* 5* (Dispersed). *Worldwide:* Total users in all countries: 27,937,610 (as L1: 27,936,530; as L2: 1,080).

10,000 to 99,999

Persian, Iranian [pes] *Users:* 89,200 in Tajikistan (2017). *Status:* 5* (Dispersed). *Worldwide:* Total users in all countries: 78,623,350 (as L1: 57,192,350; as L2: 21,431,000).

Kyrgyz [kir] *Users:* 71,400 in Tajikistan (2017). *Ethnic population:* 152,000 (2017). *Status:* 5* (Dispersed). *Worldwide:* Total users in all countries: 5,154,500.

Russian Sign Language [rsl] *Users:* 47,000 in Tajikistan (2021 DBS/DOOR/SIL). *Estimated* 28,000–56,000 signing deaf, assuming 0.3%–0.6% of total population. *Status:* 5 (Developing). *Worldwide:* Total users in all countries: 909,000.

Russian [rus] *Users:* 44,600 in Tajikistan (2017). *Ethnic population:* 303,000 (2017). *Status:* 3 (Wider communication). *De facto national working language*, no longer language of record in higher levels of government. *Worldwide:* Total users in all countries: 254,997,130 (as L1: 146,954,150; as L2: 107,987,980).

Shughni [sgh] *Users:* 40,000 in Tajikistan (1975 SIL). 20,000 Shugan, 1,500–2,000 Oroshor, 15,000 Rushan. *Ethnic population:* 73,800 (1990 A. Kibrik). Includes 50,000 Shughni, 2,000 Oroshor, 18,000 Rushan, 800 Khufi, 3,000 Bartang. *Status:* 6a (Vigorous). *Worldwide:* Total users in all countries: 80,000.

Pashto, Southern [pbt] *Users:* 32,400 in Tajikistan (2017). *Status:* 5* (Dispersed). *Worldwide:* Total users in all countries: 19,871,430 (as L1: 16,391,430; as L2: 3,480,000).

Wakhi [wbl] *Users:* 15,000 in Tajikistan (2012 M. Merritt). *Ethnic population:* 20,000 (1990 A. Kibrik). *Status:* 6b (Threatened). *Worldwide:* Total users in all countries: 52,200.

Turkmen [tuk] *Users:* 13,700 in Tajikistan (2012 UNSD). *Ethnic population:* 32,500 (2017). *Status:* Unestablished. *Worldwide:* Total users in all countries: 6,562,260.

Yagnobi [yai] *Users:* 12,000 (2004 SIL). *Status:* 6b* (Threatened).

1,000 to 9,999

Yazghulami [yah] *Users:* 9,000 (2010). Ethnic population: 13,000. *Status:* 6b (Threatened).

Armenian, Western [hyw] *Users:* 8,900 in Tajikistan (2017). *Status:* Unestablished. *Worldwide:* Total users in all countries: 1,582,290 (as L1: 1,576,490; as L2: 5,800).

Erzya [myv] *Users:* 8,800 in Tajikistan (2017). *Status:* Unestablished. *Worldwide:* Total users in all countries: 245,780.

Bashkort [bak] *Users:* 8,200 in Tajikistan (2003 J. Leclerc). *Status:* Unestablished. *Worldwide:* Total users in all countries: 1,180,940.

Aimaq [aiq] *Users:* 7,900 in Tajikistan (2017). *Status:* Unestablished. *Worldwide:* Total users in all countries: 1,891,900.

Belarusian [bel] *Users:* 6,430 in Tajikistan (2017). *Status:* Unestablished. *Worldwide:* Total users in all countries: 3,766,550 (as L1: 1,376,550; as L2: 2,390,000).

Balochi, Western [bgn] *Users:* 6,350 in Tajikistan (2017). *Status:* Unestablished. *Worldwide:* Total users in all countries: 2,283,350.

Azerbaijani, North [azj] *Users:* 5,640 in Tajikistan (2017). *Status:* Unestablished. *Worldwide:* Total users in all countries: 9,110,020.

Tatar [tat] *Users:* 4,200 in Tajikistan (2012 UNSD). Ethnic population: 107,000 (2017). *Status:* Unestablished. *Worldwide:* Total users in all countries: 5,127,970 (as L1: 4,317,970; as L2: 810,000).

Chuvash [chv] *Users:* 3,940 in Tajikistan (2017). *Status:* Unestablished. *Worldwide:* Total users in all countries: 738,150.

Lak [lbe] *Users:* 2,100 in Tajikistan (2017). *Status:* Unestablished. *Worldwide:* Total users in all countries: 100,000.

Hazaragi [haz] *Users:* 1,700 in Tajikistan (2022 Joshua Project), based on ethnicity. *Status:* Unestablished. *Worldwide:* Total users in all countries: 3,660,000.

Parya [paq] *Users:* 1,600 in Tajikistan (2017). *Status:* 6a* (Vigorous). *Worldwide:* Total users in all countries: 2,600.

Georgian [kat] *Users:* 1,400 in Tajikistan (2017). *Status:* Unestablished. *Worldwide:* Total users in all countries: 3,878,780 (as L1: 3,724,240; as L2: 154,540).

Arabic, Tajiki Spoken [abh] *Users:* 1,300 in Tajikistan (2017). *Status:* 6b* (Threatened). *Worldwide:* Total users in all countries: 17,300.

Turkish [tur] *Users:* 1,000 in Tajikistan (2017). *Status:* Unestablished. *Worldwide:* Total users in all countries: 90,028,000 (as L1: 84,010,500; as L2: 6,017,500).

100 to 999

Ishkashimi [[isk](#)] *Users:* 500 in Tajikistan (1990 A. Kibrik). *Status:* 6a (Vigorous). *Worldwide:*
Total users in all countries: 3,000.

Languages by Status

In this section the languages of Tajikistan are listed in order of their status within the country as represented by their level on the EGIDs scale (Lewis and Simons 2010). The language entries are reduced to just the information elements that are relevant to assessing the EGIDS level: population, status, language use, language development, and writing.

1 (National)

Tajik [tgk] *Users:* 6,380,000 in Tajikistan (2012 UNSD). Ethnic population: 7,520,000 (2017). *Status:* Statutory national language (1994, Constitution, Article 2). *Lg Use:* Also use Northern Uzbek [uzn]. Also use Russian [rus]. *Lg Dev:* Literature. Radio. Bible: 1992–1999. *Writing:* Arabic script, Naskh variant, used until 1928. Cyrillic script, used since 1940, primary usage. Hebrew script, used by Bukharan Jews in Uzbekistan. Latin script, used between 1928–1940. *Worldwide:* Total users in all countries: 8,287,420.

3 (Wider communication)

Russian [rus] *Users:* 44,600 in Tajikistan (2017). Ethnic population: 303,000 (2017). *Status:* De facto national working language, no longer language of record in higher levels of government. *Lg Use:* Used as L2 by Russian Sign Language [rsl], Tajik [tgk], Yagnobi [yai]. *Writing:* Braille script. Cyrillic script, primary usage. *Worldwide:* Total users in all countries: 254,997,130 (as L1: 146,954,150; as L2: 107,987,980).

5 (Developing)

Russian Sign Language [rsl] *Users:* 47,000 in Tajikistan (2021 DBS/DOOR/SIL). Estimated 28,000–56,000 signing deaf, assuming 0.3%–0.6% of total population. *Lg Use:* Vigorous. Used by all. Some also use Russian [rus]. Deaf people born before about 2000 learned Russian in school. Some also use Tajik [tgk]. Deaf people born after about 2000 have been taught Tajik in schools, and the government currently promotes Tajik over Russian. *Worldwide:* Total users in all countries: 909,000.

5 (Dispersed)

Kyrgyz [kir] *Users:* 71,400 in Tajikistan (2017). Ethnic population: 152,000 (2017). *Writing:* Arabic script, Naskh variant, used in China. Cyrillic script, not used in Afghanistan. Latin script, used in Turkey. *Worldwide:* Total users in all countries: 5,154,500.

Pashto, Southern [pbt] *Users:* 32,400 in Tajikistan (2017). *Writing:* Arabic script, Naskh variant, used in print. Arabic script, Nastaliq variant, used in handwriting and for book titles. *Worldwide:* Total users in all countries: 19,871,430 (as L1: 16,391,430; as L2: 3,480,000).

Persian, Iranian [pes] *Users:* 89,200 in Tajikistan (2017). *Writing:* Arabic script, Naskh variant. Arabic script, Nastaliq variant, used in Iran. *Worldwide:* Total users in all countries: 78,623,350 (as L1: 57,192,350; as L2: 21,431,000).

Uzbek, Northern [uzn] *Users:* 1,060,000 in Tajikistan (2017). Ethnic population: 1,230,000 (2017). *Lg Use:* Used as L2 by Tajik [tgk]. *Writing:* Arabic script, Naskh variant, symbolic use in Uzbekistan, sometimes used in China. Braille script. Cyrillic script, official usage in Uzbekistan between 1940–1992, continued widespread use in Uzbekistan and China. Latin script, official usage in Uzbekistan since 1992. *Worldwide:* Total users in all countries: 27,937,610 (as L1: 27,936,530; as L2: 1,080).

6a (Vigorous)

Ishkashimi [isk] *Users:* 500 in Tajikistan (1990 A. Kibrik). *Lg Use:* Ishkashimi is very vital. This is because the fathers very much insist on their children acquiring Ishkashimi as their first language when they grow up. Home, village. Used by all. Positive attitudes. *Writing:* Unwritten. *Worldwide:* Total users in all countries: 3,000.

Parya [paq] *Users:* 1,600 in Tajikistan (2017). *Lg Use:* Despite small population, it is spoken more vigorously amongst themselves than larger language groups. Uzbek and Tajiki men who marry Parya women learn Parya and become assimilated into the community. Home. Also use Tajik [tgk]. *Writing:* Unwritten. *Worldwide:* Total users in all countries: 2,600.

Shughni [sgh] *Users:* 40,000 in Tajikistan (1975 SIL). 20,000 Shugan, 1,500–2,000 Oroshor, 15,000 Rushan. Ethnic population: 73,800 (1990 A. Kibrik). Includes 50,000 Shugni, 2,000 Oroshor, 18,000 Rushan, 800 Khufi, 3,000 Bartang. *Lg Use:* Home. Used by all. Positive attitudes. Also use Tajik [tgk], as a literary language. Used as L2 by Wakhi [wbl]. *Lg Dev:* Literature. Dictionary. Texts. Bible portions: 2015. *Writing:* Arabic script, Naskh variant, used in Afghanistan. Cyrillic script, used in Tajikistan. Latin script, used in Tajikistan. *Worldwide:* Total users in all countries: 80,000.

6b (Threatened)

Arabic, Tajiki Spoken [abh] *Users:* 1,300 in Tajikistan (2017). *Lg Use:* Only a few speak Arabic. Shifting to Tajiki [tgk]. No diglossia with Modern Standard Arabic [arb]. Also use Tajik [tgk]. *Writing:* Arabic script, Naskh variant. *Worldwide:* Total users in all countries: 17,300.

Wakhi [wbl] *Users:* 15,000 in Tajikistan (2012 M. Merritt). Ethnic population: 20,000 (1990 A. Kibrik). *Lg Use:* Some domains. Some of all ages. Positive attitudes. Many also use Shughni [sgh]. Also use Tajik [tgk], as literary language. *Writing:* Arabic script, Naskh variant, used in Afghanistan. Arabic script, Nastaliq variant, used in handwriting and for book titles. Cyrillic script, used in Tajikistan. Latin script, used in Pakistan, preferred script, not firmly established. *Worldwide:* Total users in all countries: 52,200.

Yagnobi [yai] *Users:* 12,000 (2004 SIL). *Lg Use:* Also use Russian [rus], as a literary language.

Also use Tajik [tgk], as a literary language. *Lg Dev:* Grammar. *Writing:* Cyrillic script.

Yazghulami [yah] *Users:* 9,000 (2010). Ethnic population: 13,000. *Lg Use:* Home. Some of all ages. Positive attitudes. Also use Tajik [tgk]. *Lg Dev:* Dictionary. Texts. *Writing:* Cyrillic script. Latin script.

Unestablished

Aimaq [aiq] *Users:* 7,900 in Tajikistan (2017). *Writing:* Arabic script, Naskh variant. *Worldwide:* Total users in all countries: 1,891,900.

Armenian, Western [hyw] *Users:* 8,900 in Tajikistan (2017). *Writing:* Armenian script. *Worldwide:* Total users in all countries: 1,582,290 (as L1: 1,576,490; as L2: 5,800).

Azerbaijani, North [azj] *Users:* 5,640 in Tajikistan (2017). *Writing:* Arabic script, Naskh variant, used until 1920s. Braille script. Cyrillic script, official usage in Dagestan. Latin script, official usage in Azerbaijan since 1992. *Worldwide:* Total users in all countries: 9,110,020.

Balochi, Western [bgn] *Users:* 6,350 in Tajikistan (2017). *Writing:* Arabic script, Naskh variant. Arabic script, Nastaliq variant, primary usage. Cyrillic script, used in Turkmenistan. *Worldwide:* Total users in all countries: 2,283,350.

Bashkort [bak] *Users:* 8,200 in Tajikistan (2003 J. Leclerc). *Writing:* Arabic script, Naskh variant, no longer in use. Cyrillic script. *Worldwide:* Total users in all countries: 1,180,940.

Belarusian [bel] *Users:* 6,430 in Tajikistan (2017). *Writing:* Arabic script, Naskh variant, developed in the 16th century. Braille script, no longer in use. Cyrillic script, official usage. Latin script, official usage 1941–1944, modern occasional usage. *Worldwide:* Total users in all countries: 3,766,550 (as L1: 1,376,550; as L2: 2,390,000).

Chuvash [chv] *Users:* 3,940 in Tajikistan (2017). *Writing:* Cyrillic script. *Worldwide:* Total users in all countries: 738,150.

Erzya [myv] *Users:* 8,800 in Tajikistan (2017). *Writing:* Cyrillic script. *Worldwide:* Total users in all countries: 245,780.

Georgian [kat] *Users:* 1,400 in Tajikistan (2017). *Writing:* Braille script. Georgian (Mkhedruli and Mtavruli) script, primary usage. Khutsuri (Asomtavruli and Nuskhuri) script, no longer in use. *Worldwide:* Total users in all countries: 3,878,780 (as L1: 3,724,240; as L2: 154,540).

Hazaragi [haz] *Users:* 1,700 in Tajikistan (2022 Joshua Project), based on ethnicity. *Writing:* Arabic script, Naskh variant. Arabic script, Nastaliq variant, used in handwriting and for book titles. *Worldwide:* Total users in all countries: 3,660,000.

Lak [lbe] *Users:* 2,100 in Tajikistan (2017). *Writing:* Arabic script, Naskh variant, used until 1928. Cyrillic script. Latin script, used between 1928-1938. *Worldwide:* Total users in all countries: 100,000.

Tatar [tat] *Users:* 4,200 in Tajikistan (2012 UNSD). Ethnic population: 107,000 (2017). *Writing:*

Arabic script, Naskh variant, used prior to 1928. Cyrillic script. Latin script. *Worldwide*: Total users in all countries: 5,127,970 (as L1: 4,317,970; as L2: 810,000).

Turkish [tur] *Users*: 1,000 in Tajikistan (2017). *Writing*: Arabic script, Naskh variant, no longer in use. Braille script. Cyrillic script, used in Bulgaria. Greek script, no longer in use. Latin script, used since 1928 or 1929, primary usage. *Worldwide*: Total users in all countries: 90,028,000 (as L1: 84,010,500; as L2: 6,017,500).

Turkmen [tuk] *Users*: 13,700 in Tajikistan (2012 UNSD). Ethnic population: 32,500 (2017). *Writing*: Arabic script, Naskh variant, used in Afghanistan and Iran. Cyrillic script, used in Turkmenistan, not used in Afghanistan. Latin script, official usage in Turkmenistan since 1991. *Worldwide*: Total users in all countries: 6,562,260.

Languages by Region

This index gives an alphabetical listing of the top-level administrative subdivisions within Tajikistan. Under the name of each region is a list of the language communities that are located within its area.

Dushanbe

Armenian, Western [hyw], 10

Persian, Iranian [pes], 13

Khatlon

Arabic, Tajiki Spoken [abh], 10

Armenian, Western [hyw], 10

Kazakh [kaz], 0

Pashto, Southern [pbt], 13

Persian, Iranian [pes], 13

Turkmen [tuk], 15

Uzbek, Northern [uzn], 16

Kuhistoni Badakhshon

Ishkashimi [isk], 12

Kyrgyz [kir], 12

Shughni [sgh], 14

Tatar [tat], 15

Wakhi [wbl], 16

Yazghulami [yah], 16

Regions of Republican Subordination

Kyrgyz [kir], 12

Parya [paq], 12

Uzbek, Northern [uzn], 16

Sughd

Armenian, Western [hyw], 10

Kazakh [kaz], 0

Uzbek, Northern [uzn], 16

Yagnobi [yai], 16

Languages by Family

This index gives an alphabetical listing of the linguistic classifications used for the established languages of Tajikistan. The entries in this index represent the full path in the linguistic family tree from the highest level grouping down to the lowest. All the languages listed in the same entry are members of the same lowest-level subgroup. The referenced page contains the main entry that describes the language.

Afro-Asiatic, Semitic, Central, South, Arabic

Arabic, Tajiki Spoken [abh], [10](#)

Indo-European, Balto-Slavic, Slavic, East

Russian [rus], [13](#)

Indo-European, Indo-Iranian, Indo-Aryan, Intermediate Divisions, Western, Unclassified

Parya [paq], [12](#)

Indo-European, Indo-Iranian, Iranian, Eastern, Northeastern

Yagnobi [yai], [16](#)

Indo-European, Indo-Iranian, Iranian, Eastern, Southeastern, Pamir

Ishkashimi [isk], [12](#)

Wakhi [wbl], [16](#)

Indo-European, Indo-Iranian, Iranian, Eastern, Southeastern, Pamir, Shugni-Yazgulami

Shughni [sgh], [14](#)

Yazghulami [yah], [16](#)

Indo-European, Indo-Iranian, Iranian, Eastern, Southeastern, Pashto

Pashto, Southern [pbt], [13](#)

Indo-European, Indo-Iranian, Iranian, Western, Southwestern, Persian

Persian, Iranian [pes], [13](#)

Tajik [tgk], [14](#)

Sign language, Deaf community sign language

Russian Sign Language [rsl], [14](#)

Turkic, Eastern

Uzbek, Northern [uzn], [16](#)

Turkic, Western, Aralo-Caspian

Kyrgyz [kir], [12](#)

Language Code Index

This index gives an alphabetical listing of all 28 three-letter codes that are used in this work to uniquely identify languages. The referenced page contains the main entry that describes the language. All codes listed are part of the ISO 639-3 standard; see <http://www.sil.org/iso639-3/>.

abh	Arabic, Tajiki Spoken, 10	paq	Parya, 12
aiq	Aimaq, 10	pbt	Pashto, Southern, 13
azj	Azerbaijani, North, 10	pes	Persian, Iranian, 13
bak	Bashkort, 11	rsl	Russian Sign Language, 14
bel	Belarusian, 11	rus	Russian, 13
bgn	Balochi, Western, 10	sgh	Shughni, 14
chv	Chuvash, 11	tat	Tatar, 15
haz	Hazaragi, 12	tgk	Tajik, 14
hyw	Armenian, Western, 10	tuk	Turkmen, 15
isk	Ishkashimi, 12	tur	Turkish, 15
kat	Georgian, 11	uzn	Uzbek, Northern, 16
kir	Kyrgyz, 12	wbl	Wakhi, 16
lbe	Lak, 12	yah	Yazghulami, 16
myv	Erzya, 11	yai	Yagnobi, 16

Language Name Index

This index lists every name that appears in the language listings as a primary or alternate name of a language or dialect. The following abbreviations are used in the index entries: *alt.* ‘alternate name for’; *alt. dial.* ‘alternate name for a dialect of’; *dial.* ‘primary name for a dialect of’; *pej. alt.* ‘pejorative alternate name for’; and *pej. alt. dial.* ‘pejorative alternate name for a dialect of’. Each index entry resolves to the primary name for the language with which the indexed name is associated, followed by square brackets containing the unique three-letter language code from ISO 639-3. The referenced page contains the main entry that describes the language. If the language appears on a map, the entry for the primary name also lists page numbers for the maps on which the language occurs.

- Afghana-Yi Nasfurush**, *alt.* Parya [paq], 12
Afghana-Yi Siyarui, *alt.* Parya [paq], 12
Arabi, *alt.* Arabic, Tajiki Spoken [abh], 10
Arabic, Tajiki Spoken [abh], 10, 18
see Arabic, Tajiki Spoken [abh], 10
Western Armenian, *see* Armenian, Western [hyw], 10
North Azerbaijani, *see* Azerbaijani, North [azj], 10
Western Balochi, *see* Balochi, Western [bgn], 10
Bartang, *alt. dial.* Shughni [sgh], 14
Bartangi, *dial.* Shughni [sgh], 14
Bukhara Arabic, *alt.* Arabic, Tajiki Spoken [abh], 10
Buxara Arabic, *alt.* Arabic, Tajiki Spoken [abh], 10
Central Asian, *alt.* Arabic, Tajiki Spoken [abh], 10
Central Wakhi, *dial.* Wakhi [wbl], 16
Changgars, *alt.* Parya [paq], 12
Chuf, *alt. dial.* Shughni [sgh], 14
Eastern Wakhi, *dial.* Wakhi [wbl], 16
Eastern Yagnobi, *dial.* Yagnobi [yai], 16
Galcha, *pej. alt.* Tajik [tgk], 14
Guhjali, *alt.* Wakhi [wbl], 16
Iazgulem, *alt.* Yazghulami [yah], 16
Ishkashim, *alt.* Ishkashimi [isk], 12
Ishkashimi [isk], 12, 18
Jugari, *alt.* Arabic, Tajiki Spoken [abh], 10
Khik, *alt.* Wakhi [wbl], 16
Khikwar, *alt.* Wakhi [wbl], 16
Khuf, *alt. dial.* Shughni [sgh], 14
Khufi, *dial.* Shughni [sgh], 14
Khugni, *alt. dial.* Shughni [sgh], 14
Khugnone, *alt.* Shughni [sgh], 14
Kyrgyz [kir], 12, 18
Laghmani, *alt.* Parya [paq], 12
Lower Yazghulami, *dial.* Yazghulami [yah], 16
Oroshani, *alt. dial.* Shughni [sgh], 14
Oroshor, *dial.* Shughni [sgh], 14
o‘zbek, *alt.* Uzbek, Northern [uzn], 16
o‘zbek tili, *alt.* Uzbek, Northern [uzn], 16
o‘zbekcha, *alt.* Uzbek, Northern [uzn], 16
Parya [paq], 12, 18
Pashto, Southern [pbt], 13, 18
see Pashto, Southern [pbt], 13
Pbharya, *alt.* Parya [paq], 12
Persian, *alt.* Persian, Iranian [pes], 13
Persian, Iranian [pes], 13
see Persian, Iranian [pes], 13
Roshani, *alt. dial.* Shughni [sgh], 14
Roshorvi, *alt. dial.* Shughni [sgh], 14

- RSL**, *alt.* Russian Sign Language [rsl], 14
- Rushan**, *alt. dial.* Shughni [sgh], 14
- Rushani**, *dial.* Shughni [sgh], 14
- Rusija**, *alt.* Russian [rus], 13
- Russian** [rus], 13
- Russian Sign Language** [rsl], 14, 0
- Shighni**, *alt. dial.* Shughni [sgh], 14
- Shugan**, *alt. dial.* Shughni [sgh], 14
- Shughni** [sgh], 14, 18
dial. Shughni [sgh], 14
- Shugnan**, *alt. dial.* Shughni [sgh], 14
- Shugnan-Rushan**, *alt.* Shughni [sgh], 14
- Soghdi**, *alt.* Yagnobi [yai], 16
- Tadzhik**, *alt.* Tajik [tgk], 14
- Tajik** [tgk], 14, 18
- Tajik Sign Language**, *alt.* Russian Sign Language [rsl], 14
- Tajiki**, *alt.* Tajik [tgk], 14
- Tajiki Arabic**, *alt.* Arabic, Tajiki Spoken [abh], 10
- Tajiki Persian**, *alt.* Tajik [tgk], 14
- TSL**, *alt.* Russian Sign Language [rsl], 14
- Upper Yazghulami**, *dial.* Yazghulami [yah], 16
- Uzbek, Northern** [uzn], 16, 18
see Uzbek, Northern [uzn], 16
- Vakhan**, *alt.* Wakhi [wbl], 16
- Wakhani**, *alt.* Wakhi [wbl], 16
- Wakhi** [wbl], 16, 18
- Wakhigi**, *alt.* Wakhi [wbl], 16
- Western Farsi**, *alt.* Persian, Iranian [pes], 13
- Western Wakhi**, *dial.* Wakhi [wbl], 16
- Western Yagnobi**, *dial.* Yagnobi [yai], 16
- Yaghnabi**, *alt.* Yagnobi [yai], 16
- Yaghnobi**, *alt.* Yagnobi [yai], 16
- Yaghnubi**, *alt.* Yagnobi [yai], 16
- Yagnabi**, *alt.* Yagnobi [yai], 16
- Yagnob**, *alt.* Yagnobi [yai], 16
- Yagnobi** [yai], 16, 18
- Yagnubi**, *alt.* Yagnobi [yai], 16
- Yazghulami** [yah], 16, 18
- Yazghulomi**, *alt.* Yazghulami [yah], 16
- Yazgulam**, *alt.* Yazghulami [yah], 16
- Yazgulami**, *alt.* Yazghulami [yah], 16
- Yazgulyam**, *alt.* Yazghulami [yah], 16
- Yazgulyami**, *alt.* Yazghulami [yah], 16
- Yuzdomi zavég**, *alt.* Yazghulami [yah], 16
- Zǵamiǵai**, *alt.* Yazghulami [yah], 16
- Zgyamigyai**, *alt.* Yazghulami [yah], 16

Bibliography

Akiner, S. 1983. *Islamic peoples of the Soviet Union*. London: Kegan Paul International.

Comrie, B. (ed.). 1987. *The world's major languages*. London: Routledge.

Kibrik, A. E. 1991. The problem of endangered languages in the USSR. In R. H. Robins and E. M. Uhlenbeck (eds.), *Endangered languages*, pp. 257–273. Oxford: Berg Publishers.

Lewis, M. P. and G. F. Simons. 2010. Assessing endangerment: Expanding Fishman's GIDS. *Revue Roumaine de Linguistique* 55(2):103–120. http://www.lingv.ro/RRL_2_2010_art01Lewis.pdf. Accessed 11 January 2011.

Sebeok, T. A. (ed.). 1963. *Soviet and East European linguistics*. Current Trends in Linguistics 1. The Hague: Mouton.