

Introduction to Tajiki

Tajiki is a member of the Iranian family of languages, which includes Persian, Kurdish, Soghdian, and many other languages of Iran, Afghanistan, Tajikistan, and surrounding countries. (*Tajiki* is the name of the language, *Tajik* the name of the people.) Tajiki is one of the three varieties or dialects of Persian, the others being Farsi (Iran) and Dari (Afghanistan). The Iranian languages in turn make up part of one branch (the Indo-Iranian languages) of the Indo-European family of languages, which includes Greek, Latin, the Germanic languages (including English), the Slavic languages (including Russian), the Celtic languages, the Indian languages (such as Sanskrit and Hindi), Armenian, and others. That is, all of these languages are modern descendants of a much earlier language called Proto-Indo-European, which was never written down but which linguists can more or less reconstruct. Some Tajiki words are very close to the corresponding words (cognates) in English and the other Indo-European languages:

Tajiki	English	Latin	Greek
modar	mother	mater	matér
padar	father	pater	patér
bar-	bear (v.)	fer-	pher-

Other words have changed so much in the different languages that they are quite dissimilar—for example, Tajiki *zan* ‘woman, wife’ is cognate with English *queen*, Greek *gyné*, Russian *zhená*, and Irish *ban* (as in *banshee*, ‘woman of the fairies’).

Proto-Indo-European seems most likely to have been spoken in the steppes north of the Black Sea at the time it started spreading. The beginning of the breakup of Proto-Indo-European into its daughter languages cannot be dated exactly, but rough estimates can be made on the basis of archeological evidence and the absence of certain terms relating to agriculture and technology common to all the branches of Indo-European languages. Proto-Indo-European probably broke up no later than 5,000 to 6,000 years ago. After the language spread, it broke up into different dialects that increasingly diverged from each other as their speakers lost contact with each other. This gave rise to separate languages that in turn spread and diverged over the millennia to develop into the various languages of each of the branches of Indo-European family.

The people who spoke the ancestral form of the Indo-Iranian languages (Proto-Indo-Iranian) spread into the region south



of the Black, Caspian, and Aral seas well before 1,000 BC. By the first millennium BC, Proto-Indo-Iranian had already separated out into distinct Indian and Iranian languages. The Iranian languages at this time (the Old Iranian languages) were spoken from the north shore of the Black Sea to west and northwest of China (making up the dominant language of the Scythians, a nomadic culture dominating the Eurasian steppes and including peoples speaking early Turkic and Mongolic languages), and further south in the Iranian plateau to the shores of the Persian Gulf. Numerous peoples spoke Old Iranian languages, including

the Persians proper along the Persian Gulf, the Medes in the northwestern Iranian plateau, and various Central Asian peoples: Soghdians, Bactrians, Khwarezmians, Parthians, and others. The major linguistic division at this time was between the Western Iranian (Old Persian and Median) and Eastern Iranian languages (the others); the Western Iranian languages are further classified as Northwestern (Median) and Southwestern (Old Persian).



The Medes unified the region of the Iranian plateau and expanded to found an empire extending throughout much of modern-day Iran; its Persian successor states, the Persian empire of the Achaemenids (648-330 BC), the Parthian empire (170 BC-226 AD), and the Sassanid empire (226-650 AD), ensured the cultural and linguistic unity of the Western Iranian languages. The Eastern Iranian languages, on the other hand, were spoken in a much more fragmented region in which Soghdian served as a lingua franca (a common language of trade and administration) that coexisted with local languages which came to diverge much more from each other over this period than the Western Iranian languages did; the descendants of these languages (which are spoken throughout Afghanistan and Tajikistan) are much more distinct from Tajiki (and from each other) than Tajiki is from Farsi and Dari. Moreover, while the Persian empires were officially Zoroastrian (such religions as Judaism, Nestorian Christianity, Roman and Greek paganism, and Buddhism were usually tolerated, though Catholic Christianity was often persecuted under the Sassanids as the official religion of Persia's major enemy, Rome), Central Asia at the time was a cultural and religious stew in which Buddhism, Manichaeism, Nestorian Christianity, and Zoroastrianism coexisted. Indeed, each religious tradition used a distinct script to set itself apart from the others; however, these scripts were not limited to particular languages but rather adapted to the various local languages when that religion's scriptures were translated. Culturally, Central Asia was a network of city-states based on agriculture and trade linked by the various trade routes of the Silk Road surrounded by a sea of nomadic herding peoples, in which local leaders rarely became powerful enough to conquer their enemies and were often prey to the depredations of nomadic steppe empires and conquest by

sedentary empires like Persia and China. This is the cultural background which even today shapes Tajik culture as much as Islam has.

After the conquest of Persia by the Arabs in 650, the Eastern Iranian languages went into decline as Islam was spread to Transoxiana (also known by its Arabic designation, *Mawara' an-Nahr*, "the land beyond the [Oxus] river"). The Persian of the time, written in Arabic script and enriched with a number of Soghdian loanwords, became the new common language of culture, trade, and administration throughout Central Asia and further west to the region around Nishapur; it was the form of Persian spoken in many of the great cultural centers of Central Asia: Bukhara, Samarkand, Balkh, Tashkent, and Khujand, for example. It was also the form of Persian spoken by many of the greatest Persian poets and philosophers after the Arab conquest of Persia, like Rudaki, Firdawsi, Omar Khayyam, Rumi, Attar, al-Farabi, ibn-Sina (Avicenna), and al-Ghazali. This preeminence was established by the first native Persian Islamic dynasty, the Samanids (819-999), whose capitals at Samarkand, Bukhara, and Herat are among the greatest cultural centers of Central Asia. At this time Persian was viewed as a unitary language called *Dari*, meaning "of the court"; regional differences in speech were unimportant compared to the uniformity of the written language, which has changed little since the 9th century.

However, with the rise of the Safavids in the early 16th century, this cultural unity was weakened. The Safavids, who expanded from Azerbaijan to rule Iraq, Iran, and much of Afghanistan, were staunch Shi'ites who warred constantly with the Sunni states on their borders (indeed, the dynasty was effectively ended in 1722 by the revolt of the Ghalzai, Afghan Sunnis whom the Safavids had tried to convert by force to Shi'ism).



Once the Persian cultural world was divided, the language of the two regions diverged, with that of Persia coming to be called *Farsi*; the term *Tajiki* for the eastern dialects is a 20th century innovation. (The older term for the language, *Darii Kobuli*, meaning "Persian of Kabul," is still used for the Persian spoken in Afghanistan.) Outside of language, the religious and cultural differences between the Safavids and Transoxiana accentuated the Central Asian orientation of Tajik culture. However, there is a tendency among scholars to call the common language *Farsi* regardless of national boundaries, much as American English is called English rather than American.

Finally, a century of Russian domination encouraged the spread of Western culture as refracted through Russian culture and Soviet ideology (including a significant influx of Russian loanwords, a major portion of which is common Greco-Latin terminology borrowed into Russian); in particular, the Czarist and Soviet policies towards minority nations inculcated western ideals of nationalism

that are common to Central Asia but differ significantly from the view of nation and state in much of the rest of the Islamic world. In addition, in accordance with Soviet policy for minority languages, a new written standard of Tajiki was promulgated that was based heavily upon colloquial speech, which led to the elimination of many Arabic words common in literary Persian.

Besides Russian loanwords (many of which are being replaced by non-Russian words), Tajiki has also borrowed many words from Arabic and some (though far fewer) from the Turkic languages. Arabic words are especially common in the literary and intellectual levels of the language, very much as Greek and Latin words are especially frequent in these realms in English, but they are also quite common in the most commonplace levels of speech—thus, the Arabic word *ba'd* 'after' is at least as common as the native Persian word *pas*, and both prepositions are fully acceptable in speaking. Similarly, compound verbs based on *tavallud* (Arabic for 'birth') are as commonly used as the native Persian verb *zodan*. Turkic words are most common in vocabulary concerning livestock, herding, and steppe life, but even terms for body parts and kinship terms have been borrowed from the Turkic languages. In return Arabic and the Turkic languages have borrowed heavily from Persian. (It is said that throughout the Islamic world Arabic is the language of religion, Persian the language of administration and poetry, and Turkic the language of the military. More important for the beginning student of Tajiki, Arabic is also the language of grammar.) While most of these words will be unfamiliar to most speakers of English, there are a few Arabic words that have been borrowed into both Persian and English: *sifr* 'zero' goes with both *cipher* and *zero*, and *sharbat* 'juice' with *sherbet* and *sorbet*, for example. There are also a number of Persian words that have been borrowed into English (often through Arabic). Thus, *noranj* 'orange' goes with *orange* and *pojoma* goes with *pajama* (in Tajiki it refers to the light undertrousers worn by both sexes, from *poi* 'leg' and *joma* 'dress'). Similarly, *vizier* is the English form of Tajiki *vazir*, and by a circuitous route *divan* comes from the Persian word *divan* (Tajiki *devon*) for a document house or ministry. The Persian word was also used for the hall in which an administrator would hold public audiences, and then this was applied to the sofa where he sat. (The Persian word also referred to a collection of poems, a meaning occasionally also found in the West, as in Goethe's *Divan of West and East*, or *West-Östlicher Divan*.)

