Manuscripta Orientalia

International Journal for Oriental Manuscript Research
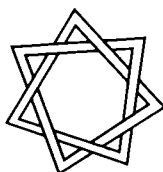
THESA
St. Petersburg

P. B. Lurje

ARABOSOGDICA: PLACE-NAMES IN TRANSOXIANA AS WRITTEN IN ARABIC SCRIPT*

Among all of the non-Persian Iranian toponyms of Mawarannahr that have come down to us thanks to medieval sources (some 1,000 items [1]), the vast majority (80—90%) are those mentioned in Islamic Arabic, Persian and Turkic sources. This multilingual group is united by certain general characteristics — a single cultural environment, a great deal of mutual influence, and, most important, a single tradition of writing. The transmission of Sogdian glosses through the Arabic script is the topic of the present article.

The Persian and Turkic varieties of writing are far better suited to Sogdian phonetics than their Arabic precursor. The Arabic alphabet with all additional Persian letters (*p*, *č*, *ž*, *g* and *β*, Arabic *f* with three dots — *fā' 'ajamī*, *fā' muthallath*, found in some early New Persian and Turkic texts, Chorasmian glosses) can correctly convey virtually all Sogdian consonants. But the Arabic alphabet itself could have served the needs of Sogdian no worse than could the Buddhist (Sogdian-Aramaic), Christian (Nestorian) or Manichaean writing systems (cf. the various difficulties in conveying Sogdian labials with these writing systems). Moreover, one should bear in mind that Persian texts, especially the early ones that are most important for us, often lack all of the necessary diacritics, making them the functional equivalent of Arabic-alphabet texts. Additionally, some Arab authors — especially al-Sam'ānī (henceforth,

Sam.) — often indicate "special" Persian letters in the headings of their dictionary entries [2].

The preceding comments on the means Arabic script possesses to convey Sogdian sounds would hold if the authors of the sources used here [3] had written coherent Sogdian texts rather than individual glosses in a language they appear not to have known at all [4]. Taking into account the slips of the pen that inevitably creep into manuscripts, many words are difficult to recognize. For example, a village in the land of the Qarluqs on the way to Barskhān (a city on Lake Issyk-Kul), Qudāma (206) calls كرمينكث [5], Gardīzī — كومبر كث and *ḤA* — كرمينكث [6]. Only by juxtaposing all of these forms can we reconstruct the name with reasonable surety as كرميركث*, Sogdian *krm'yr+knδh/Karmīr-ka"θ/* 'Red Town' [7].

However, in other cases variant spellings reflect not so much errors and carelessness in transcription, but various ways of conveying the sounds [8]. For example, the well-known settlement and palace of the Bukhār-khudāts to the west of Bukhara, usually referred to in the archaeological literature as *Varakhsha*, is given in the most varied forms in Islamic sources: *Afraxšī*, *Abraxša*, *Awraxšā*, *Faraxša* [9]. This allows one to reconstruct the name as Sogd. *(Ϡ)βraxšē*. This article treats precisely such means of conveying Sogdian sounds in Arabic script.

1. Consonants [10]

1.1 Back. We know that Sogdian had three back consonant phonemes (*k*, *γ*, *x*) and one allophone (*g<k* in the *ng<nk* position). Early Sogdian had a *x^y* sound, later contracted to *xu*. The Arabic alphabet has letters for the first three sounds, but lacks a letter for the fourth (*g*), so the *ng* group was conveyed in two ways: with *-nj* (cf. Arab. *šaṭranj* < Pers. *šaṭrang*) or (more rarely) with *nk*. The latter spelling was more typical of Persian usage. Cf. the numerous place-names in *-sank/-sanj*, the majority of which go back to Sogd. *sang* 'stone'; some of them, however, proba-

bly go back to Sogd. Budd. *snk'/sanga/* 'monastic community' (<Old Ind. *saṃ-gha-*). Moreover, one should take into account that *x* could be distorted in the copying process to *ḥ* or *j*, *ghayn* to *'ayn*, *k > l* (and *vice versa*). We note especially that Sogdian *k* could also be given as an ordinary *k* or emphatic *q*. While *q* was used primarily in Turkic glosses, it is also found in Sogdian names. For example, the name of the Bukharan settlement *Saqmatīn* in Narshakhī (II, XXII, pp. 8, 52) apparently goes back to Sogd. *(')sk('/y)+myδn*, Old Iran. **uska-maiθana-*, 'Upper Settlement' [11]; a very

---

late (15th century) name of a ploughed field near Samarqand, *mazra'a-yi rīmīq* [12], apparently means 'dirty field', where *-rīmīq* is derived from Sogd. *rym* 'dirt', in the same way as *''pyk* 'watery' < *''p-*. *Xᵛ* was usually given as *x* with a following *u ḥarakat*, sometimes (especially before long vowels) as *xw*. For example, the village and region not far from Bukhara (IḤq, 486, Iṣt., 310, Sam., 195v, Yāqūt, II, 427 [13]) — *Xurmīθan* from Sogd. *xwr + myδn* 'Settlement of the Sun' [14]; the old name of Kokand — *Xwākand* (*passim*, Iṣt. — *Xākand*, late medieval *Xūkand*), probably from Sogd. **xwā(k)* 'trade' [15] + *knδh* 'town'. In one instance we find a confusion of *x/γ* [16] in the name of the village *Yaγnī* near Nakhshab (Sam., 601r — 601v, Yāqūt, III, 1022, Sam. D., 11, 86), which goes back to Sogd. Man. *yxnyy* 'food set aside' (lit. 'icy', 'frozen'), cf. *Yaγnāb*, not attested in available [17] Arabic-script sources (while 'ice' in Yaγnobi is actually *ixn*). Cf. also the confusion of *q/γ* in clusters: *Xaryān-rūd* (Iṣt., 309—10), but *Xarqān-rūd* in Narshakhī (II, Frye, No. 7, cons.).

1.2. Sogdian had one affricative — *č* with a variant as *j* (in *nj*). Moreover, N. Sims-Williams reconstructs the sound *c* (*ts*) [18] on the basis of *ts* given as *c* (*č*) in Sogdian Christian texts. Arabic has only the affricative *j*, which did not convey Sogd. *č* in all cases. In conveying well-known names [19], *č* was transcribed with *ṣ* (*Ṣaγāniyān*, Pers. *Čaγāniyān*, Sogd. attributive *cγ'ny*,). This was typical also of Arabic spellings of loans from Persian and other languages (*ṣaidana* < Ind. *čandan* 'sandal'; *ṣirāj* < NPers., Sogd. *čarāγ*, etc.). Another possible variant was the writing of *č* with *š*, cf. *Šāš* < *Čāč*. The same phenomenon is observed in other parts of Iran, for example, the Arabic name of the Caspian Sea town *Čālūs* as *Šālūs*. Finally, *j* was used for *č* in the names of small objects found only in isolated passages in geographical works. We have, for example, *Jīnanj(a)kaθ* for *Čīnanjkaθ*, the name of two cities, one (smaller) in *Čāč*, the other (bigger) in Chinese Turkestan. The latter is mentioned in a Sogdian sale-contract (*cynncknδ-*) [20]. V. Minorsky was the first to translate this name as "Chinese town" [21]. We also note that we lack credible examples of place-names with *c* (*ts*) (including those on the territory of Northern Bactria and Khorezm).

1.3 It would appear that Arabic script has all the means to convey Sogdian dental phonemes — *t*, *θ*, *δ* and the allophone *d*. But it would be naïve to expect that copyists did not confuse *t* and *θ* or *d* and *δ*, which differ from each other by a single dot. This makes it impossible to establish the crucial distribution of the forms *-kat/-kaθ*; some early sources use one of the two forms, while other copies use various forms. Sometimes *t* was voiced to *d* (*δ*). One recalls the nickname of the Arab governor of Central Asia, Sa'īd 'Abd al-'Azīz, a lover of luxury — *Khudhayna* 'Lady, empress' (Sogd. *xwt(')ynh /xutēna/*). In some cases this spelling may have been the result of processes occurring in Sogdian itself [22], but in other cases stems from Persian influence on Arabic spelling [23]. *T* in Sogdian names was usually given as ordinary *t*, although in some cases emphatic *ṭ* was used, sometimes under the influence of popular etymology. For example, Yāqūt and Sam. (III, 538) [24] call one village near Bukhara *Ṭaγāmī*, deriving the name from the Arabic *Ṭaγām* 'lower elements, the mob', although in actuality it comes from Sogd. *ty'm(+'k)* 'fording place' [25]. By the same token, Sogdian *s* could be given with emphatic *ṣ*, but we found no cases of *z* or *ḍ* used for Sogdian *z* or *d(δ)* [26].

1.4. Arabic is least effective for conveying the labials. Sogdian had the following labial phonemes: *p*, *f*, *β* [27], *w* [28] and the allophone *b* (in *mb* < *np*, *zb* < *zp*). Arabic writing has signs only for *w*, *f* and *b*; the remaining Sogdian sounds in this group were variously conveyed, and only a juxtaposition of spellings can confirm a certain reading. *B* or *f* were used to convey *p*. For example, the region on the lower Syr-Darya and its tributary, Arys, was variously called *Bārāb*, *Fārāb* and *Fāryāb*. The original form of the name can be reconstructed as *Pār'āp*, which appears to mean 'irrigated lands' (Pers. *fāryābī* — *idem*) or 'opposite bank of the river' (Avest. *pāra-*, Shugn. *por*, etc. 'opposite bank' [29]) + Sogd. *''p* [30]. *β* was conveyed with *w*, *b* and *f* (with equal frequency, apparently). The latter usage is explained by the presence of a letter *β* (*fā' 'ajamī, fā' muthallath*) in early Persian, Turkic and Khorezmian texts (see above the numerous variants of Sogdian *(∃)βraxšē* — *Varakhsha*). Only by juxtaposing variants can we succeed in reconstructing a place-name; otherwise, we must rely on intuition. The sound *w* could sometimes become *u* (and vice versa). Thus, the formant derived from the Sogd. *w'δ('k)* 'place' gives either *-waδ* or *-ūδ*. Cf. *γaδāwaδ* (IḤq, 493, Iṣt., 316, Yāqūt, III, 776, vowelled *γuδāwaδ* in Sam., 406v) — *maḥalla* (quarter) and gate in Samarqand, which apparently means "Place of thieves" [31], on the one hand, and *Arfūd* (Sam., 206v, Yāqūt, I, 209 — *al-Arfūd*, possible the name with the spelling variant *Arqud*, Narshahī, IV, 13) — a village in Central Soγd (the first part of the name is unclear). This was, however, typical not only of Arabic spellings, but of the Sogdian language in general — cf. Mugh *kštwt*, but *ʼftm'wt* — modern Kištut and Falmowut.

1.5. The sibilants *s*, *š*, *z* had parallels in the Arabic alphabet, unlike the sound *ž*. There were no difficulties in conveying *s* and *z*; in most cases, the sound *š* was given as *š*, sometimes as *s* or *ṣ*. This is the case with the place-name *Ṣāγarj* (IḤq, 500, Sam., 285v, Yāqūt, II, 11, the last two authors indicate the possibility of an *s/ṣ* alternation in the name), which N. Sims-Williams [32] sees as going back to Sogd. *š'wyrc-* 'relating to the Black Mountain' or the name of the canal *Ṣāfarī-kām* in Iṣt., 310, which corresponds to *Šāfur-kām, Šāpūr-kām* in most other sources. This phenomenon occurs with the names of other large centres: *Kiš/Kaš, Naxšab/Naṣaf*. Cf. the same within the Iranian plateau — Arab. *Naiṣābur* for the Pers. *Nēšābūr*, MPers. *Nēw-šāhbuhr*. Examples of place-names with *ž* (all of minimal credibility) show that this sound was given with *z* (the village *Zāz* near Ishtikhan in Sam., 266v — from the Pers. *žāž* 'nettles'(?); a well-known canal and region near Bukhara — *Farāwī/āz* — from the Sogd. *pr'wyz /parāwēž/* 'flow, irrigate'(?)) or with *j* (*Γajdawān* for modern *Γižduwān*); cf. also *Nūzābāδ/Nūjābāδ* (Sam., Yāqūt, s.v.) — both forms indicate Persian *nōž-ābāδ* 'Pine-tree-place' [33].

1.6. The sound *r* was almost always given as *r*, but in Ferghana, Chach and Ustrushana place-names, sometimes as *l*. This notable circumstance, however, is likely less a consequence of Arabic writing than Sogdian dialect particularities; as such, it deserves independent discussion.

1.7. The sonorous *y* has an Arabic equivalent. Like the half-vowel *w*, *y* could be inserted before a corresponding vowel or when an etymological *y* was lost, changing the vowel. One of the cities in Ispijāb, which IḤq (510, 516) and Iṣt. (337) give as *Buδaxkaθ*, is found in Muq. (263) as

Iδaxkaθ and in Yāqūt (IV, 1014) and Sam. (598v) as *Yaδaxkaθ* (possibly as the result of a copyist's error).

1.8. Finally, the nasal consonants *n* and *m*. The Arabic alphabet has letters for both, and little variation is encountered here. We do not dwell here on the distribution of the forms *kaθ/kand*, which requires additional study. In one (dubious) case, we find an *m/b* alternation. One of the large centres to the west of Samarqand, *(A)rabinjan* (*passim*), is

called *Ramījan* [34] in a fifteenth-century Samarqand land document (Sam. D., 10, 33—5).

1.9. Correspondences between consonant phonemes in Sogdian and the Arabic alphabet are given in the *Scheme* at the end of the present article. To the right are listed Sogdian consonants; to the left, Arabic letters. A solid line indicates an ordinary correspondence; a dotted line, rarer variants. Sogdian allophones are given in brackets.

2. Vowels

Overall, Arabic writing has the same resources for conveying vowels as the three main types of Sogdian writing. Moreover, Arabic has a comparatively strict system of diacritical vowelling symbols (*harakāt*), which allows one to indicate short vowels. But for foreign glosses, vowelling (obligatory only for quotes from the Qur'ān) is chaotic [35], often incorrect [36], and lacks special symbols for the "*majhūl*" vowels *ē*, *ō* and *schwa* (*ə*), which are not found in Arabic. Among the general features of Arabic vocalism we also note the comparative "stinginess" of the *matres lectionis*; spellings such as (*')'y* for a single vowel /*i, ī, e, ē*/ (typical of Sogdian) are unthinkable. Finally, one should bear in mind that the classical Arabic and medieval Arabic linguistic tradition did not tolerate double consonants at the beginning of a word, two consonants after a long vowel, or three consonants in a row.

2.1. Long *ā* is given as *alif*. In final position, this could be written as *yā'* (*alif maqṣūra*), cf. the frequent spelling of *Buxārā* as *Buxārī* (read as *Buxārā* in the nominative). In initial position, the double *alif* (Ī) is a rarity found only occasionally in Sam. and Yāqūt (*Āfuran* (Sam., 15r, Yāqūt, I, 64), a village one *farsakh* away from Nakhshab < '*p+βr'n* 'Bringing water'(?)), in *ḤA* (4a, No. 3.32; 9b, No. 6.23) — *Āwāza-i Paikand* for the modern Lake Qaraqul to the south-west of Bukhara, from Sogd. *''w'zh* 'lake' and in later texts.

2.2. *Ī* is always given as *y* ('*y* in initial position).

2.3. *Ē* is usually given as *ī*. At the end of a word, *ē* (from **-a-ka*) is usually transcribed as *h* (*-a*, cf. Yaγn. *-a* from Sogd. *-ē*), more rarely *ā* or *ī*, possibly *-iya*, sometimes the final *k* reappears. We note that for the numerous names in *-dīza*, the Arabic adjective (*nisba*) was *-dīzakī*, where a *k* reappears. The nature of the final sound is unclear for the Arabic spelling *-iya* or *ī* (*Dabūsiya/Dabūsī*, *Kušāniya/Kušānī*, *Karmīniya/Karmīna* [37] — all three are large centres between Samarqand and Bukhara). Medial *ē* was sometimes vowelled with a *fatḥa*; it should be read as the diphthong *ai*. Cf., for example, the Bukharan canal *Baikān* (Narshahī, XIII, tr. Frye, 32) [38] — possibly from *βēk+ān* 'Outer'. Moreover, long *ē* could be given as *alif* (the so-called *imāla*). See 1.5 above for the example *Farawī/āz*, also *Isbārī/Isbīrī* in Sam. (290v) and Yāqūt (III, 36). In some cases such variants could indicate an *ā*. For example, the village *Farruxšāδ* in Ṭab. (II, 1540, 112) corresponds to *Farruxšīδ*, one of the gates of Samarqand (IḤq, 493, Iṣṭ., 316) [39], and should possibly be understood as the MPers. *Farrux*ˢ*-šād* 'joy of Farrukh' [40].

2.4. *Ū* (like the case with *ī*) was given in initial position as '*w*, and in medial position as *w*.

2.5. *Ō* was conveyed like *ū*. We found no variants in our sources (unlike the case with *ē*).

2.6. Short vowels (*a, i* and *u*) were conveyed with *alif* at the beginning of a word and with Ø in medial position. The case with the majority of final vowels in Sogdian is not entirely clear. In some place-names we find a confusion between *w*/Ø and *y*/Ø. This, for example, is the case with the name of a district to the north of Samarqand — *Būzmājan* (*passim*) and *Buzmājan*, a district in Kesh (IḤq, 502), *Šī/irbadūn* (Bukh. D., 1, 441), a lake near Bukhara. One should also note that the metathesis *CuCC > C(u)CuC*, typical of Sogdian, has not Arabic equivalent. The name for Sogdiana — *S/Ṣuyd* (only read as *Suyud* after Tīmūr), a small city near Kesh *Xušminjkaθ* (Sam., Yāqūt) — in honour of the Sogdian deity *xšwm*; cf. n. 14 above. Vowels could be labialised after labial consonants. Cf. the numerous cities with the name *Bunjīkaθ* [41]. Meanwhile, the best-known of them was called *pnčyknδ(h)* in Sogdian; also Arab. *Buttam(ān)* for Sogd. *pyttm'n* [42] — a mountainous land to the south of Zarafshan. It is possible that the numerous place-names in *-ūn* go back to the formant *-ān*, widespread in the Iranian world, with the *ān>ūn* shift common to many Iranian languages. We are not aware of a single reliable Sogdian name in *-ū* [43].

2.7. There are numerous examples of vanishing initial *alif* (vowelled *a-* or *i-*), meant to convey Sogdian reduced *ə-*. Cf., for example, *(A)bārkaθ*, a village to the north of Samarqand, the capital of the above-mentioned district of *Būzmājan* (a form with the *a-* *passim*, with *bār-* Ṭab., III, 82—3, Iṣṭ., 332, Sam., 59r, Yāqūt, I, 464). This name may be derived from the OIran. **upari* 'Upper' or (more likely) the OIran. **pāri* 'opposite side of the river' [44], as the settlement was located directly to the north of Zarafshan on a passage from Samarqand that lies on the south bank of the river.

2.8. There is some variation in the indication of diphthongs (long and short). For example, Sogd. *š'w* 'black' is almost always given as *šāw-* [45] (we have five toponyms in *šāw-*). Otherwise, we find Sogdian *z'y* 'land' (i) as *zā* (cf. *Zāmīn*, a well-known city in Ustrushana, *passim*, modern *Zaamin*); Abdunabiev (*Ura-Tiube*, 7) proposed for this name the etymology 'earthen (*z'y*) + suburb (Yaγn. *men(a)*, *main*)'; a plot of land near Tashkent (Sam. D., 12, 53) *šīrzā* 'good (*šyr*) land'; (ii) as *za* (cf. *Fāriza* — a castle and *maḥalla* in Bukhara in Sam., 415v, and Yāqūt, III, 855; numerous etymologies are possible for the first part); (iii) as *zī*, cf. the village, long absent in the sources (before the nineteenth century?). It appears, for example, in the memoirs of Sadriddin Ayni and on Russian maps, but apparently of ancient origin, *Wayānzī*, to the northeast of Bukhara — from the Sogd. *βγ'n+z'y* 'land of the gods(?)'; (iv) finally, as *-ziya*, cf. *Wāyziya*, one of the three regions of Ferghana in Muq. (262), possibly from *β'γ+z'y* 'land of gardens'.

3. Persian mediation

3.1. As was noted above (see n. 2), the majority of our authors did not know Sogdian. But they all knew Persian to some degree; hence, many names and elements that in Sogdian form differed little from Persian, were redone in Persian fashion.

3.2. Sogdian *šāw* 'black' was always given correctly (cf. above, 2.8). The Persian correlative *siyāh* is quite different. But Sogdian *(ə)spēt* 'white' is clearly similar to Pers. *saf/p/bēd/δ* (note final voiced). As a result, nearly all names with this element were conveyed in Persianized form: *Sabīdmāša* (Iṣṭ., 309, IḤq, 484) near Bukhara, *Sabīδyūk(ī)* (Sam., 290r, Yāqūt, III, 36), in the same region (in the second part, Sogd. *ywk*, Pers. *yōk* 'frog(?)' [46]), *Saba/īδmūn* (Sam., 284v, Yāqūt, III, 31 (*Sabaδyūn*), Bukh. D., 1, 244). The same voicing took place with Sogd. *rwt* 'river', which in place-names became the Persianized form *rōδ, rōd*. Cf. *Rūδak* (Sam., 262r, Yāqūt, II, 833; Rūdakī's native land), *Jādrūd*, the old name of Magyan-darya to the east of Penjikent (Sam. D., 10, 207), possibly came from *c'δr+rwt* 'lower river', *Asrūd* (IḤq, 501, Iṣṭ., 324), a river in Kesh, etc. [47] A well-known section of Samarqand is called by Sam., 489r [48] *Mātrīt* (the second part apparently from the Sogd. *ryt* 'face'). But later (until quite recently, in fact) [49], it was called *Mātirīd* with a voiced

final consonant, yet another possible example of phonetic Persianization.

3.3. In a number of cases, we find translations (not transcriptions) of names. The well-known crossing between Termez and Kesh is called by all sources the "Iron gates": Arab authors — *Bāb al-ḥadīd*; Persian *Dar(b)-i Āhanīn*, Turkic *Temir qapïy*; the *History of the T'ang Dynasty* [50] also calls it the "Iron gates". In the Sogdian original, this name was apparently */Ʒspnēnē-δβar/*. In the same fashion, the gate and surrounding area of Samarqand, known to the Arabs as *Ra's al-Qanṭara*, Pers. *Sar-i pul*, Uzbek. *Köprik-boši*, could once have been called **Yitk-sar-*. Ya'qūbī (294) calls one of the cities of Ustrushana *Ḥiṣnak* [51]; this is apparently the well-known city of *Dizak* (*passim*, today *Jizzax*), literally 'small fort'. The author translated the root of the name (Pers. *dī/ē/iz*, Sogd. *δyz* 'fort' = Arab. *Ḥiṣn*), but left the suffix unchanged. Among such names, the most notable are *al-Qariya al-ḥadītha* [52]/*Deh-i naw* [53]/*Yaŋïkänt* [54], a small town in the Syr-Darya delta, where all names translate as 'new town'. But Muq. (263) gives the form *Deh Nūjkat* for this town. It is easy to see in this name contamination between the Pers. *Deh-i Naw* and Sogd. */Nawč+ka"θ/* 'new city' [55], which was apparently the original name.

* * *

Research in all areas of onomastics is of crucial importance when sources of other linguistic material are limited. One can cite as an example the extensive research on Old Iranian personal names, which have substantially enriched our knowledge of ancient Iranian vocabulary, allowing us, for example, to identify the Median language, throw light on the road taken by the Iranians to the Iranian plateau, etc. M. Mayrhofer, a leading specialist on Old Iranian onomastics, in the subtitle to one of his articles noted the "central" role of onomastics in ancient Iranian linguistics [56]. We remind readers that the overwhelming majority of ancient Iranian personal names are attested in other languages — Elamic, Akkadian, Greek, Aramaic, etc. In other words, the majority of ancient *Nomina Propria* are at the same time *Nebenüberlieferungen*, where the names are often heavily distorted.

We encounter the same phenomenon in the place-names of medieval Transoxiana. Autochthonous sources [57] list some 100 place-names, roughly as many as are found in Chinese sources; several dozen geographic names can be found in Old, Middle Persian, Parthian, and Armenian works. But the 1,000 place-names that have come down to us in Arabic-script sources far outweigh all of the above-mentioned groups. This is why we consider it indispensable to investigate how these place-names were conveyed in Arabic writing. As we have seen above, in many cases Sogdian toponyms from Mawarannahr are not merely conveyed through a different language, but are, so to speak, *Doppelnebenüberlieferungen* — filtered through two languages with Persian as an intermediary.

I n d e x *

''fr'n	'sbyry	bx'r'	dr(b) ''hnyn	gd'y	kš'ny(h)
''w'zh' pyknd	'srwd	bx'ry	dry'm	gd'ych	kštwt
'b'rkθ	'wrxš'	byk'n	δrym	yjdw'n	kwmbrkθ
'brxšh	b'b 'l-ḥdyd	bzm'jn	dzk	ḥsnk	kwprk b'šy
'δxkθ	b'r'b	č'č	f'r'b	j'drwd	m'trb
'frxšy	b'rkθ	č'lws	f'ry'b	jn	m'tryd
'ftm'wt	bδxkθ	čy'ny'n	f'rzh	jynnjkθ	m'tryt
'l-qryh 'l-ḥdyθh	bdy'n'	**čy'ny**	fr'w'z	jzx	m'trb
'rbnjn	(jwyb'r)bk'r	**čynncknδ-**	fr'wyz	kBrmBr'w	nsf
'rfwd	bmjkθ	dbwsy(h)	frxš'δ	krmyn(y)h	nwj'b'δ
'rqwd	bnjykθ	dh byd	frxšh	**krmyrkθ*	nwjkθ
'rmw	bṭm'n	dh nw	frxšyδ	kṣ	nwkθ
'sb'ry	bwzm'jn	dh nwjkt	yδ'wδ	kš	nwyjkθ

nwykt	*rbnjn*	*sbydm'šh*	*š'š*	*x'knd*	*yγn'b(?)*
nwz'b'δ	*rmyjn*	*sbyδmwn*	**š'wγrc**	*xrγ'nrwδ*	*yγny*
nxšb		*sbyry*	*šrbdwn*	*xrq'nrwd*	*yngγknt*
nyṣ'bwr	*rwδk*	*ṣγ'ny'n*	*šwy*	*xrmk*	*z'z*
nyš'bwr	*(mzr'h-y)*	*ṣγd*	*šyrbdwn*	*xrmydyzh*	*z'myn*
pdy'nh	*rγmyq*	*ṣγd*	*šyrz'*	*xrmyθn*	*zrnjry*
pncyknδh	*s'fryk'm*	*sqmtyn*	*ṭγ'my*	*xšmnjkθ*	*zrnkry*
pyttm'n	*s'γrj*	*sr'pl*	*tmyr q'pγ*	*xw'knd*	
r'mtyn	*sb'ry*	*š'frk'm*	*vsp š't xwsrw*	*xwknd*	
r'myθn	*sbdmwn*	*š'pwrk'm*	*wγ'nzy*	*xžw'n (*xžδ'n)*	
r's'l-qnṭrh	*sbyδγwk(y)*	*š'lws*	*w'γzyh*	*yδxkθ*	

Scheme

**Correspondences between consonant phonemes in Sogdian
and the Arabic alphabet**
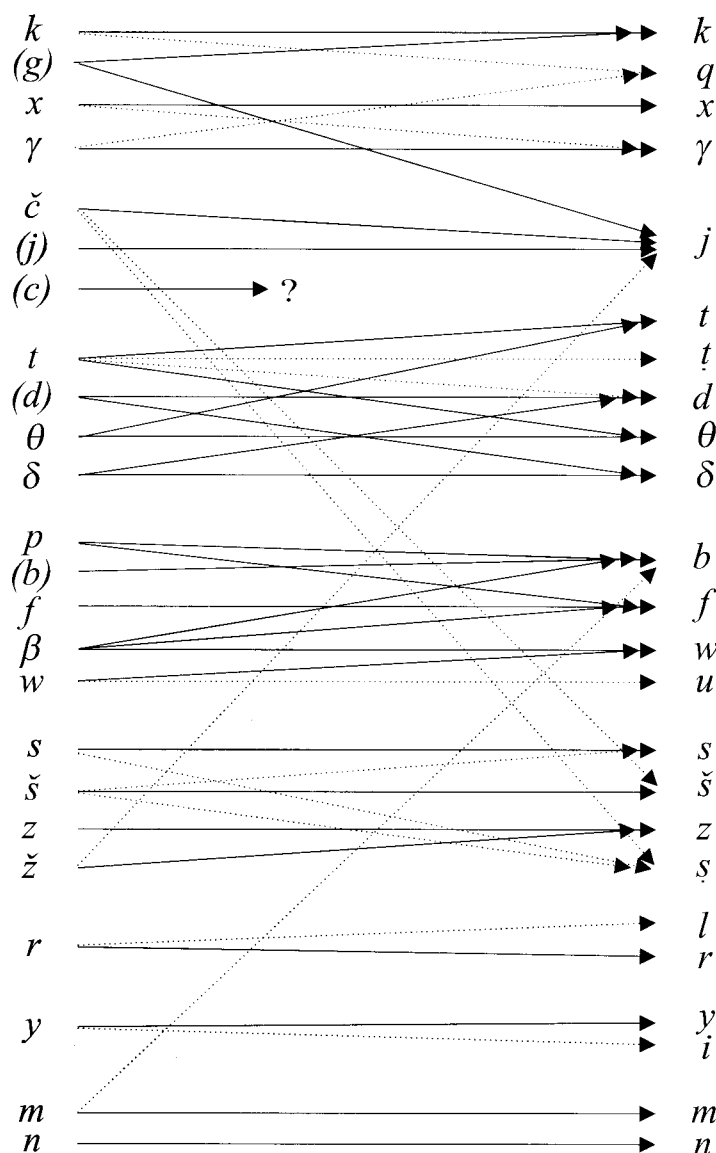
Abbreviations

Abdunabiev, *Ura-Tiube* — A. Abdunabiev, *Lingvisticheskiĭ analiz toponimii regiona Ura-Tiube* (A Linguistic Analysis of the Ura-Tiube Region Toponymics), avtoreferat kandidatskoĭ dissertatsii (author's summary of the PhD dissertation) (Dushanbe, 1992).

Barthold, *Turkestan* — V. V. Barthold, *Turkestan down to the Mongol Invasion* (London, 1958). — Gibb MS, V (reprint).

BGA — *Bibliotheka Geographorum Arabicorum*, edidit M. J. De Goeje, Lugdini-Batavorum, apud E. J. Brill.

Bukh. D. — O. D. Chekhovich, *Bukharskie dokumenty XIV veka* (Fourteenth-century Bukharan Documents) (Tashkent, 1965).

Eilers, Demavend — W. Eilers, "Der Name Demavend", *Archiv Orientální*, XXII, pp. 267—374; XXIV, pp. 183—224 (Zusatznoten); XXXVII, pp. 416—48 (Index zu...).

EIr — *Encyclopaedia Iranica*.

Gershevich, *GMS* — I. Gershevich, *A Grammar of Manichaean Sogdian* (Oxford, 1961).

Gharib, *Dictionary* — B. Gharib, *Sogdian Dictionary* (Tehran, 1994).

Gibb MS — E. J. W. Gibb Memorial Series (London).

ḤA — *Ḥudūd al-'Ālam* (The Regions of the World), translated and explained by V. Minorsky (London, 1937). — Gibb MS, New Series, XI.

Henning, Loan-words — W. B. Henning, "Sogdian loan-words in New Persian", *BSOS* (1939) [=Selected Papers, I].

Iṣṭ. — Abū Isḥāq Ibrāhīm Muḥammad al-Fārisī al-Iṣṭakhrī, *Masālik al-mamālik* (*Viae Regnorum*). — *BGA*, I, 1927.

IḤq — Abū al-Qāsim Ibn Ḥawqal al-Naṣībī, *Kitāb ṣūrat al-arḍ*, ed. J. H. Kramers (Leiden, 1939).

Markwart, *Wehrot* — J. Markwart, *Wehrot und Arang* (Leiden, 1938).

Muq. — Muḥammad Ibn Aḥmad Maqdisī (Muqaddasī), *Kitāb aḥsan al-taqāsīm fī ma'rifat al-aqālīm*. — *BGA*, III, 1906.

Kamaliddinov, *Istoricheskaia geografiia* — Sh. S. Kamaliddinov, *Istoricheskaia geografiia Iuzhnogo Sogda i Tokharistana po arabskim istochnikam* (An Historical Geography of the Southern Sogd and Tokharistan by the Arab Sources Accounts) (Tashkent, 1996).

Khromov, Toponimika Maverannakhra — A. L. Khromov, "O strukturnykh osobennostiakh iranskoĭ toponimii Maverannakhra v period IX—XII vv." ("On the structural features of the Iranian toponymics of the Mawarannahr in the 9th — 12th centuries"), *Vostochnaia filologiia*, fasc. III (Dushanbe, 1974).

Narshahī — *The History of Bukhara*, translated from a Persian abridgement of the Arabic original by Narshakhī by Richard N. Frye (Cambridge, 1954).

Qudāma — Qudāma ibn Jaf'ar, *Excepta e Kitāb al-kharaj*. — *BGA*, VI, 1889.

Rosenberg, Sogdica — Frederic Rosenberg, "Sogdica I", *Prace Lingwistyczne Ofiarowane Janowi Baudouinowi De Courtenay* (Krakow, 1921), pp. 94—6.

Savina, *Slovar'* — V. I. Savina, *Slovar' geograficheskikh terminov i drugikh slov, formiruiushchikh toponimiiu Irana* (The Dictionary of Geographical Terms and Other Words Forming the Toponimics of Iran) (Moscow, 1971).

Sam. — *The Kitāb al-ansāb of 'Abd al-Karīm as-Sam'ānī*, with an introduction by D. S. Margoliuth. — Gibb MS, XX, 1912.

Sam. D. — *Samarkandskie dokumenty XV—XVI vv.* (Samarqand Documents of the 15th — 16th Centuries), a critical edition with facsimile, translation, and commentary by O. D. Chekhovich (Moscow, 1974).

Sims-Williams, *Upper Indus* — N. Sims-Williams, *Sogdian and Other Iranian Inscriptions of the Upper Indus* (London, 1992).

Sims-Williams, Sogdian — N. Sims-Williams, "Sogdian", *Compendium Linguarum Iranicarum* (Wiesbaden, 1989).

Smirnova, Karta — O. I. Smirnova, "Karta verkhoviĭ Zeravshana po mugskim dokumentam" ("A map of Upper Zarafshan by the documents from Mugh"), XXV International Congress of Orientalists. The USSR delegation's papers (Moscow, 1960).

Steblin-Kamensky, Tri slova — I. M. Steblin-Kamenskiĭ, "Tri sogdiĭskikh slova" ("Three Sogdian words"), *The 50th Anniversary of the Excavations of Ancient Penjikent*, theses of the papers delivered at a scientific conference (Penjikent, 1997).

Steblin-Kamensky, *ESVJa* — I. M. Steblin-Kamenskiĭ, *Ėtimologicheskiĭ slovar' vakhanskogo iazyka* (An Etymological Dictionary of the Vakhi Language) (St. Petersburg, 1999).

Ṭab. — Abū Jaf'ar Muḥammad aṭ-Ṭabarī, *Ta'rīkh umam wa al-mulūk = Annales quod scripsit*, ed. M. J. de Goeje (Lugdini-Batavorum, 1879—1901); cf. also *Istoriia at-Tabari* (Ta'rīkh by al-Ṭabarī), trans. from the Arabic into Russian by V. I. Beliaev, with additions by O. G. Bolshakov and A. B. Khalidov (Tashkent, 1987).

Tafazzoli, Tree words — A. Tafazzolī, "Three Sogdian words in the Kitab al-Huruf", *Bulletin of Iranian Cultural Foundation*, I/2 (1973).

Ya'qūbī — Aḥmad Ibn Abī Ya'qūb al-Ya'qūbī, *Kitāb al-buldān*. — *BGA*, VII, 1892.

Yāqūt — Yāqūt ibn 'Abd-Allāh, *Kitāb mu'jam al-buldān* (Jakut's geographisches Woerterbuch), ed. F. Wuestenfeld (Leipzig, 1866—1873).

Yoshida—Moriyasu, Sale-Contract — Yutaka Yoshida, Takao Moriyasu, "A Sogdian sale-contract of a female slave from the period of Gaochang kingdom under the rule of Qu clan", *Studies on the Inner Asian Languages*, IV (Kobe, 1988).

Notes

1. This includes toponyms fully or partially intelligible on the basis of the Sogdian, Bactrian and Khorezmian languages, Turkic-Sogdian composites (with the exception of oikonyms in *-kand/qant*), as well as geographic names from the region that are not entirely clear, but obviously Iranian in origin.

2. For example, *Padyāna* (Sam., 69 r = Yāqūt, I, 527 — *Badyānā*) with a symbol for *p* that is lacking in Arabic conveyed the Sogd. /paδyāné/, lit. 'of the foot, podal' (> 'lower'), a settlement near Bukhara. Cf. also *Xižwān* (Sam., 198 r = Yāqūt, II, 440 — *Xizwān*). Read *Xižδān* and cf. Sogd. *γzδ'n* 'Grave?'

3. These are primarily Arab geographical works published by De Goeje in *BGA* (we note that the Kramers edition was used for Ibn Ḥawqal), dictionaries by Sam. and Yāqūt, the Persian anonymous geography *Ḥudūd al-'Ālam*, historical sources beginning with Ṭabarī, later land documents from Samarqand and Bukhara, and other sources.

4. In addition to his native Khorezmian, Arabic, Persian, and rudiments of Syriac and Greek, Bīrūnī probably knew some Sogdian: the anonymous author of *ḤA*, for example, indicates that the Bukharan Manichaeans were called *niγōšak*, a correct rendering of the Sogdian *nγ'wš'k*, etc. — 'a listener, *auditor*, the lowest rung in the Manichaean hierarchy'. Two Sogdian (Bukharan) phrases have come down to us in extremely distorted form through Narshakhī (see Rosenberg, Sogdica, pp. 94—6; in seemingly independent fashion, through W. B. Henning *apud* Narshakhī, comm., p. 136). Three (correct!) Sogdian glosses are found in the works of the philosopher al-Fārābī, see Tafazzoli, Three words.

5. The *hamza* is used for a letter with no dots.

6. See *ḤA*, comm., p. 92.

7. The distribution of the forms -*kand* and -*kaθ* should be the subject of a separate article.

8. The closely related question of transmitting Middle Persian loan-words in Arabic is perfectly summarized by A. Tafażżolī in *EIr*, I, pp. 231—3.

9. See Barthold, *Turkestan*, p. 115.

10. In reconstructing Sogdian phonetics, we relied mainly on Gershevich, *GMS*, Nos. 82—482.

11. Narshakhī usually gives Sogd. -*mēθan* ('settlement') (*mīθan* to the early Arab geographers) as *matīn*, cf. *Rāmatīn* (Narshahī, II, p. 8 ff.) for *Rām(m)ēthan* 'Settlement (of divinity, day of the month?) Rām' or 'Peaceful settlement' (*passim*). We will discuss the many Sogdian place-names based on the names of divinities in a separate article.

12. Sam. D., doc. 10, p. 32.

13. *Xarmaiθan*.

14. Or 'settlement (where they hold a fair on the) 11th day of the month — *xwr(-rwc)*'; hardly possible from Sogd. *x(w)r(w)m* 'earth, soil' + *myδn*, cf. the village of *Xurmak* (Sam. D., doc. 12, 59), *Xarmīdīza* (*ibid.*, doc. 18, 83).

15. Cf. Sogd. *γw(') 'k(k)r* 'trader', MPers. Man. *wh 'g* 'price', Bact. *οαυαγο* 'idem'.

16. As V. B. Henning noted, this confusion is typical of Sogdian words borrowed into Persian, see Henning, Loan-words, p. 97 [=p. 643].

17. *Yaγnāb* appears to be mentioned in *Kitāb al-kand fī dhikr 'ulamā' Samarqand* ("Sugar-book about the Scholars of Samarqand") by the 11th-century author al-Nasafī, a work not available to me. It was published not long ago in Cairo. A linguistic analysis of the Iranian glosses in this text was undertaken by the late Prof. A. Tafazzoli; V. A. Livshits was able to familiarize himself briefly with the latter's index cards during Tafazzoli's trip to St. Petersburg and told the author about the reference to *Yaγnob* in this work.

18. N. Sims-Williams, Sogdian, p. 178.

19. We mean here the names of large regions found not only in specialized geographical and historical texts, but in all other literature as well.

20. Yoshida—Moriyasu, Sale-contract, pp. 5, 6, 8.

21. V. Minorsky in *ḤA*, comm., p. 271.

22. See Gershevich, *GMS*, Nos. 268—269, also Sims-Williams, Sogdian, p. 179. One should note that this type of voicing in place-names (especially in the form *kad* < *kaθ*) was most typical of Southern Sogd (Nakhshab and Kesh).

23. For more on this, see the present article, 3.2.

24. This article is missing in the manuscript published by Margoliuth, but it can be found in the recently published *Bāb al-khalq*, Cairo, 1357—1369/1979—1992, 4 vols., which is a conspectus of Sam'ānī's work drawn up by Ibn al-Athīr (*al-Lubāb fī tahdhīb al-ansāb*), III, pp. 87—8.

25. We note here another phenomenon: in Sogdian words borrowed into New Persian, *θ* and *δ* are regularly given as *l* (see Henning, Loan-words, p. 97 [=I, p. 643]. We were, however, unable to find examples of Sogd. *δ, θ* given as *l* in medieval place-names in Mawarannahr.

26. There is, however, one river in Southern Tokharistan called *Darγām* (not to be confused with *D/Δarγam* — a canal and region near Samarqand), but it was south of the border with the region in question.

27. Also transcribed as *v*.

28. Of course, the sound *w* is described in Sogdian phonetics as a half-vowel, but for the description of its spelling with the Arabic script we place it together with the labial consonants.

29. For the latest account, see Steblin-Kamensky, *ESVJa*, pp. 283—4.

30. Similarly, Eilers, Demavend, I, 369, n. 228.

31. Cf. the semantically similar *Gadāyi*, *Gadāyča*, etc. in Iran, from *gadā* 'indigent, beggar' (see Savina, *Slovar'*, p. 50).

32. Sims-Williams, *Upper Indus*, p. 72.

33. Since all the other place-names with -*ābāδ* (Old Iran. **ā-pāta-*) are definitely Persian (see Khromov, Toponimiia Maverannakhra, p. 11), it is impossible to seek Sogd. *n(')wc* 'new' (fem.) in the latter form of the name of this town.

34. In the second part, the word *jan* ('stream, canal'), based on a toponym identified by J. Markwart (Markwart, *Wehrot*, p. 81). As far as we know, this word is not found in Sogdian texts and has not reliable etymology. Markwart (*ibid.*) suggested correcting *jan* to **xan*, a word related to the Persian *xān* 'stream'. I. M. Steblin-Kamensky (Three words, pp. 30—1) corrects (following O. I. Smirnova) *Jan* — the name of the lake where the river Sogd originates (Lake Iskanderkul) — to *Jay* and derives it from the unattested Sogdian *žā/ay* < Old Iran. **drāya-* 'sea'. One can object to the two aforementioned etymologies, which concern all hydronyms in -*jan*, that such names are numerous (more than 30) and one cannot correct them all. Of greater interest is A. Abdunabiev's suggestion (see Abdunabiev, *Ura-Tiube*, pp. 7—8) that *jan* (**/čan/?*, in modern usage — *jan, jin, jon*) should be derived from some Old Iran. form of the root **kan-* 'dig'.

35. Only Sam. and Yāqūt provide it always.

36. Sam. and Yāqūt call one of the villages near Bukhara *Zaranj/kari*, apparently derived from the Sogd. /*zirn-karē*/ 'jeweler'.

37. The second ends in something akin to Yaγn. *mēn(a)*, *main* 'quarter, area'.

38. This is apparently the correct form of *Jūybār-Bakār* in IHq (p. 484) and Ist. (p. 307).

39. The city gates usually bore the names of nearby villages.

40. Cf. the numerous names of the form *Visp-šād-xusraw* 'all-joyful Xosrow' which dotted the map of Sassanian Iran.

41. I know of four such toponyms: two in Eastern Turkestan (*ḤA*, 17a, 12.2, 18a, 15.8), the capital of Ustrushana and Penjikent to the east of Samarqand (*passim*); one should possibly include *Bumijkath*, the settlement in the Bukharan oasis (*passim*).

42. See Smirnova, Karta, p. 4.

43. With the possible exception of *Armū* (Sam. D., 17, 69, a settlement in the *tūmān* (district) *Suγud-i kalān* near Samarqand).

44. See above, 1.4 in the present article (about *Pārāb*).

45. Possible exceptions: *ṣāyarj*, see above, 1.5 in the present article; it is hardly worth deriving *Šawwī*, a city beyond Balāsāgun in Semiryechye, in Muq. (p. 264), from *š'w+k* 'black'; more likely, it should be read as **sūy* and compared with the city of *Sūyāb* (Ṭab., II, 1441, 1596) and *Su-je* in Chinese sources, the nowadays site of Ak-Beshim.

46. The Sogdian word is absent in B. Gharib's dictionary (see Gharib, *Dictionary*), but is found in fragment III of "Old letters", see Henning, Loan-words, p. 95 [=I, p. 641].

47. A. L. Khromov (see Khromov, Toponimika Maverannakhra, pp. 15—6) suggested that toponyms in *rōδ* are probably of Persian, rather than Sogdian, origin. Indeed, we failed to find toponyms with this formant with indubitably Sogdian first parts, but the word *rwt* 'river' itself is common to Sogdian.

48. The article heading (*nisba*), apparently written *al-Mātirīnī*, with the vowelling gives the utterly fantastic *wa 'ākharuhā yā' manqūṭa min fauq* (and at the end *yā'*, the last of the letters with the dots on top (sic!)); a correct description; Yāqūt (IV, 373) gives the form *Mātirab*.

49. Nowadays — *Maturit*, the latter form (with final *t* again) is due to Russian influence (similarly, another modern section in Samarqand is *Dagbit*, doubtless from Persian *Dah bēd* "Ten Willows").

50. See Kamaliddinov, *Istoricheskaia geografiia*, pp. 122—3.

51. *BGA*, VII, 294.

52. IḤq, 512. In the commentary to *ḤA* (p. 371), V. Minorsky also gives *al-Qariya al-Jadīda*, but I did not find this form in the primary sources.

53. *ḤA*, 26a=26, No. 29.

54. Late Medieval *Yanykent*.

55. Cf. the numerous Sogd. *Naw-kaθ, Nōčkaθ, Nawīkaθ, Nawīčkaθ*.

56. M. Mayrhofer, "Aus der Namenwelt Alt-Irans. Die zentrale Rolle der Namenforschung in der Linguistik des Alt-Iranischen", *Innsbrucker Beiträge zur Sprachwissenschaft*, Vortrag 3 (Innsbruck, 1971).

57. Primarily texts from the Mugh mountain, "Old letters", Sogdian inscriptions from the upper Indus valley, etc.