



SOGDIAN LANGUAGE I. DESCRIPTION

SOGDIAN LANGUAGE

i. Description

Sogdian is one of the Eastern Middle Iranian languages once spoken in Sogdiana (northern Uzbekistan and Tajikistan) before the Islamization of the area in the 10th century. Sogdians were traders along the Silk Roads and founded many diasporas along the routes, with the result that the bulk of its materials was discovered in [Turfan](#) and [Dunhuang](#) in western China. The Sogdian language was written in three scripts: Sogdian, [Manichean](#), and [Syriac](#). While only religious texts were written in Manichean and Syriac scripts, any kind of texts, both religious and secular, are recorded in Sogdian script, which was a kind of a national script, although it ultimately originated from [Aramaic](#) script.

The history of [Sogdiana](#) is largely obscure. It constituted a satrapy of the [Achaemenian](#) Empire, which was conquered by [Alexander the Great](#) in the fourth century BCE. Later it was governed by or under the influence of neighboring empires, such as the [Kushan](#) (1st to 3rd centuries CE), [Sasanian](#) (3rd to 4th centuries), Kidarite (5th century), [Hephthalite](#) (6th century), Western Turk (6th to 7th centuries), and Tang China (7th to 8th centuries). However, until Sogdiana was conquered by the Arabs in the 8th century it



enjoyed a degree of independence, during which the Sogdians played an active role as international traders along the Silk Roads between China and the West. So much so that Sogdian became a kind of lingua franca of the region and the nomadic peoples like Turks and Uighurs adopted it as their administrative language (see [SOGDIAN TRADE](#)).

Since the language of [Samarqand](#) and [Bukhara](#) of the 10th century reported by Moqaddasi is nothing but a local variety of New Persian (Yoshida, 2009a, p. 330), Sogdian had began to lose its ground by that time. However, Moqaddasi's statement that Soğd (the area between Samarqand and Bukhara) had its own language which was similar to that spoken in the suburb of Bukhara seems to indicate that at that time Persian was still restricted to the urban area (Moqaddasi, tr., p. 335). Later it gradually gave way to Persian and then to Turkish. A language called [Yaghnobi](#) still spoken in the valley of Yağnāb, a tributary of the Zarafšān River, is the only descendant of a dialect of the Sogdian language and is sometimes referred to as "Modern Sogdian," whereas the language of texts known to us is believed to represent the standard Sogdian spoken in Samarqand.

Studies of the language were began by [F. W. K. Müller](#) (1904, pp. 96-103; [Figure 1](#)), who cited two manuscripts discovered by the German Turfan Expedition (see [Turfan Expeditions](#)). At that time he referred to the language as "ein bisher nicht bekannter Pehlevi-Dialekt," although he mentions "soghdisch" later in "Nachträge und Verbesserungen" (idem, p. 111) without giving any reason. Later, in 1908 it became clear that it was [F. C. Andreas](#) who first identified the language; he compared some Sogdian [calendar](#) terms recorded by [Biruni](#) with those found in the Turfan texts (Andreas *apud* Müller, 1908, p. 3, n. 3). The first Sogdian grammar was written by Robert Gauthiot (1914-23) and [Émile Benveniste](#) (1929), who mainly referred to the Buddhist and Christian texts published by that time. This was greatly improved by [W. B. Henning](#) (1937), who based his analysis on phonologically more reliable materials written in Manichean script. The most extensive description of the grammar remains that of I. Gershevitch (1954; to be consulted with Gershevitch, 1945), which covers all aspects of the grammar, including the syntax. Since then major improvements have been made by N. Sims-Williams, who has published a series of important articles mainly on the phonology and morphology. His discoveries were incorporated into his compact description published in 1989. Sims-Williams's later discoveries have been included in Y. Yoshida's (2009a) description. Apart from useful glossaries accompanying



editions of substantial texts, three lexicons have been made public: B. Gharib (1995), Sims-Williams and Durkin-Meisterernst (2012), and Sims-Williams (2016a). Personal names found in Sogdian texts are collected in Lurje (2010).

Materials and publications. Most of the Sogdian manuscripts have been discovered in what is now western part of China, in particular from the oasis of Turfan and the so-called “library cave” of the Caves of the Thousand Buddhas in Dunhuang (see [DUNHUANG ii. Buddhist and Other Texts in Iranian Languages](#)). An important exception is the site at Mount Mug in Sogdiana proper (120 km east of Samarqand) where 70-odd documents of the early 8th century were unearthed.

Apart from the so-called [Ancient Letters](#) of the early 4th century, inscriptions discovered in Kultobe in Kazakhstan (see Sims-Williams et al., 2007; 1st-2nd century CE?) and in the Upper Indus valleys (see Sims-Williams, 1989/1992 and 2000, Yoshida, 2013; 5th century?), the bulk of Sogdian materials dates back to the 7th to early 11th centuries. The latest datable material is the inscription of 1026 CE discovered in Kyrgyzstan (Livshits, 2008, p. 379, idem, 2015, p. 290). While the Mug documents represent part of the archives of a local ruler [Dhēwaštīč](#) (ruled until 722), almost all the manuscripts from Turfan and Dunhuang are religious texts translated from the originals in Chinese (Buddhist texts), Middle Persian and Parthian (Manichean), and Syriac (Christian texts; see [CHRISTIANITY iv. Christian Literature in Middle Iranian Languages](#)).

While virtually all the texts discovered in Dunhuang have been published, a substantial number of the Turfan texts remain unpublished. The studies of the Buddhist, Manichean, and Christian texts have been surveyed, respectively, by Yoshida (2009b), Sundermann (2009), and Sims-Williams (2009). Useful catalogues and lists of the Turfan texts have also been made public (Boyce, 1960; Kudara et al., 1997; Reck, 2006; Morano, 2007; Sims-Williams, 2012; Reck, 2016). Recent substantial editions include Durkin-Meisterernst and Morano (2010), Sundermann (2012), and Sims-Williams (2014, 2015). On the Mug documents, which include letters, administrative, economic, and legal documents, see the fundamental publications by Livshits (2008) and its English version published in 2015 and Bogoljubov and Smirnova (1963). Some texts were revised by Grenet and la Vaissière (2002) and Yakubovitch (2002, 2006). Upper Indus inscriptions were read by Sims-Williams (1989/1992). Since [Hans Reichelt](#) (1931) published the Ancient Letters, some of them have been revised by Sims-Williams (2001a, Letter II) and Sims-Williams et al. (1998, Letter V).



See also Sims-Williams (2005). For several secular documents and inscriptions of various periods see Sims-Williams (1993, Ladakh), Yoshida and Moriyasu (1988, Astana, Turfan), Sims-Williams and Bi Bo (2011 and 2015, Khotan), Yoshida (2005, Xi'an), Livshits (2008/2015, Semirech'e, etc.). For the three inscriptions discovered in Mongolia, see Kljashtorny and Livshits (1971, Sevrey; 1972, Bugut), Yoshida apud Moriyasu and Ochir (1999, pp. 122-25, Bugut; pp. 225-27, Sevrey), and Yoshida, 1988 and 2011 (Karabalgasun); see also [KARABALGASUN ii. The Inscription](#).

Scripts, orthography, and basic phonology. Sogdian texts are written in three scripts: Sogdian, Manichean, and Syriac. While Manichean script is restricted to Manichean texts and Syriac to Christian, Sogdian script is a kind of national script employed for all kinds of texts, regardless of the religious affiliation of the writer. On the scripts, see Skjærvø (1996) and Yoshida (2009a, pp. 281-84). In this article, spellings in the three scripts are indicated by S[ogdian], M[anichean], and C[hristian], when distinction is necessary. Apart from them a handful of fragments written in [Brahmi](#) script are known (cf. Maue and Sims-Williams, 1991; Sims-Williams, 1996c). Sogdian script is an adaptation of the Achaemenian chancellery script deriving from Aramaic, from which resulted the phenomenon called “heterography” where several Aramaic words appeared in their original spellings but were pronounced with their Sogdian equivalents (see ARAMAIC). However, heterograms or [ideograms](#) in Sogdian are far fewer than in Middle Persian or Parthian. In the printed texts these ideograms are transliterated with capital letters, e.g., *CWRH* (cf. Aram. *šwr-h* “his neck, throat”) was pronounced as *yrīw* “self, body.” However, employment of the ideograms is not consistent at all, in that the corresponding Sogdian forms are often spelled out in different places of the same text, e.g., *CWRH* ~ *ryrywh*, *KZNH* = *m’δ* “thus,” *ZK* = *xw* (article).

Because of its long history, texts written in Sogdian script contain many historical spellings, from which those in Manichean and Syriac scripts are virtually free, e.g., *xš’yδ* “king” < OIr. **xšāyaθya*; cf. M *xš’yδ*. In the texts later than the 7th or 8th century, two styles of Sogdian script are discerned, formal and cursive (Sims-Williams, 1976, pp. 44-45). The former used to be referred to as “sūtra script,” since it was encountered mainly but not exclusively in Buddhist texts or sūtras. The cursive script was later applied to write down the Uighur language and became Uighur script, which developed further into Mongolian and Manchu scripts (Sims-Williams, 1981a, 2016b; Kara, 1996; Moriyasu, 1997). It is of some interest that the Sogdian alphabet differs from



the Uighur counterpart in that the former still includes *d* (daleth), *t* (teth), *ʿ* (ʿain), and *q* (koph), which are not employed for transcribing Sogdian sounds, while the latter disposes of them totally. However, the forms of these four letters in the Sogdian alphabet are completely different from what one expects as the natural developments from the original Aramaic letters (Livshits, 1970, 2008, pp. 298-305, 2015, pp. 227-32; Yoshida, 1995). Possibly under Chinese influence, Sogdian script began to be written vertically rather than horizontally in the latter half of the 5th century (Yoshida, 2013). In fact, all the epitaphs later than Kultobe and Upper Indus inscriptions are inscribed vertically.

Sogdian is a dead language, and the phonemic tables given below are largely based on comparative phonology and internal reconstruction. The phonetic values of certain sounds are inevitably speculative (see Table 1 and Table 2). In discussion of phonology, Sogdian forms are cited in orthographic as well as in vocalized forms; in succeeding sections transliterated forms are mainly given. For understanding of the Sogdian phonology and morphology, the so-called “rhythmic law” is of vital importance (Sims-Williams, 1981a, 1984, 1989, pp. 181-82). According to this law, the retention or loss of proto-Sogdian final vowels depends on whether they bore a stress; in its turn, the position of the stress in words of more than one syllable is determined by the quantity of those syllables: all stems in Sogdian, whether nominal or verbal, containing a long vowel are heavy and bear a stress, while a stress falls on the endings of light stems that consist only of short vowel(s). In what follows, light stems are written with a final hyphen (e.g., *wn-* [wan-] “to do”) to distinguish them from heavy stems (e.g., *wyn* [wēn] “to see”).

All three writing systems, which ultimately originate from Aramaic, are consonantal and have many features in common; they are hardly adequate for representing the vowels of Sogdian. The vowels are indicated or rather hinted at by means of three *mater lectiones*, i.e., *ʿ*, *w*, and *y*. Thus *ʿ* (initially “as well”) represents [ā], *w* [w, u/ū, o/ō] and *y* [y, i/ī, e/ē], although the short vowel [a] is not spelled out except for the initial and final positions (indicated by *ʿ* or *-ʿ*, respectively). A somewhat better situation is acquired by Syriac script, where the employment, though not consistent, of special diacritics enables the distinction between [i/ī] and [e/ē], [u/ū] and [o/ō]; sometimes word medial [a] is also indicated by two diacritic points. Unfortunately, the texts written in Brāhmī script are not many and are not of much help in this respect (Maue and Sims-Williams, 1991; Sims-Williams, 1996c).



Some pairs are cited to illustrate the vowel phonemes: *βay-* (*βγ-*) “god,” *βāy* (C *b'γ*) “garden, farm”; *fneš-* (C *fnyš-*) “to be deceived,” *fnēš* (C *fnyš*) “to deceive”; *witar-* (*wytr-*) “to go,” *wītar* (*wytr*) “he went” (3rd sg. imperfect); *but-* (*pwt-*) “Buddha,” *pūtē* (*pwt'k*) “rotten”; *roxšn-* (*rwxsšn-*) “light, bright,” *rōyn* (*rwyn*) “oil.” In Table 1, the sounds in parentheses are variants of the phoneme /a/. In addition to the simple vowels shown above, Sogdian possessed three rhotacized vowels: *ə'*, *i'*, and *u'* and a nasal vowel *ṁ*. The rhythmic law shows that a syllable containing a simple vowel followed by a rhotacized or nasal vowel is treated as heavy, e.g., *maə'γ* (*mry*) “forest,” *kaṁθ* (*knδh*) “city” in contrast to light stems such as *mə'γ-* (*mry-*) “bird,” *kī'm-* (*kyrm-*) “snake,” *pu'n-* (*pwrn-*) “full.” There are some uncertainties surrounding the vowel written as *wy*. Most of the forms in question originate from **u/wa/wā* followed by **y* or **aya*, e.g., *wyžp-* “terror” < **ubjyā-*, *xwyr* “sun” < **xwarya-*, and *xwyr* “to feed” < **xwāraya-*. The fact that in the Uighur orthography front rounded vowels *ü* and *ö* are represented by *wy* suggests that Sogdian also possesses the front rounded vowels, while the spellings in Brāhmī script does not seem to support this assumption (Sims-Williams, 1981a; idem 1996b). In any case, most forms lose the *-y-* element in later texts, e.g., C *wžp-* “terror” and C *xwr* “sun.”

In Table 2, the sounds in parentheses are allophones (voiced stops and [ŋ]) and marginal phonemes (*ts*, *l*, and *h*) mainly employed in foreign words. Complex consonants clusters do not seem to be avoided, as in *roxšn-* (*rwxsšn-*) “light (adj.),” *xēpθt* (C *xypθt*) “one’s own (pl.),” *yu'δtk-* (C *ywrδtq-*) “kidney.” Typologically unusual combinations of voiced fricative and voiceless plosive or affricate are also not uncommon, e.g., *əzpaə't* (*'zprt*) “pure,” *āδč* (“δc”) “something,” *paδk-* (*pδk-*) “law.” The second consonant is likely to be realized as a weak, voiceless plosive (Sims-Williams 1989a, p. 179).

Correspondences between the twenty-two letters of the Aramaic alphabet and Sogdian phonemes are as follows: ' (*aleph*): *a*, *ā*, *ə*; *b* (*beth*): *β*, *f*; *g* (*gimel*): *γ*; *d* (*daleth*): only in ideograms; *h/H* (*he*; only used in word-final position): *-a*, zero (cf. Sims-Williams 1981a, p. 350, e.g., *wana* (S *wnh*) “tree,” *rāθ* (S *r'δh*) “road”; feminine nouns usually end with *-h*; Wendland (1998) transcribes *-H* and takes it for an ideogram indicating feminine forms); *w* (*vau*): *w*, *u/ū*, *o/ō*; *z* (*zain*): *z*, *ž*; *x* (*cheth*): *x*, (*h*); *ṭ* (*teth*): not attested; *y* (*jod*): *y*, *i/ī*, *e/ē*; *k* (*kaph*): *k*, (*g*); *δ/L* (*lamed*): *δ*, *θ*, (*l*); *m* (*mem*): *m*; *n* (*nun*): *n*, *ṁ*; *s* (*samech*): *s*; ' (*ain*): only in ideograms; *p* (*pe*): *p*, (*b*), *f*; *c* (*tzaddi*): *č*, (*j*); *q* (*koph*): not attested; *r* (*resh*): *r*, *V'*, (*l*); *š* (*schin*): *š*; *t* (*tau*): *t*, (*d*). Note that the letter *lamed* is transliterated as *δ* when it represents Sogdian sounds [δ] and [θ]. On the historical background in



which the letter *lamed* came to stand for [δ], see Sims-Williams, 2016b. When one letter stands for more than one sound or two letters coalesce to assume one and the same form, diacritics are not uncommon among the texts later than the 8th century, e.g., *frāyāzt* (S *fr'y'zt*) “he begins” with β provided with an extra stroke, *žamnu* (*zmnw*) “time, hour” with a point below *z*, which is otherwise indistinguishable from *n* and *z* (Sims-Williams, 1978, pp. 257-58). Some of the diacritics are shared by the Uighur script and could have originated from the Uighur usage.

In the texts written in the Manichean script, new signs were introduced to transcribe β, δ, γ, and the letters *b*, *d*, *g* represent the voiced allophones of *p*, *t*, *k*, while the letters *q* and *ṭ* are interchangeable with *k* and *t*, respectively. ‘(*ain*) stands for the prothetic vowel [i], an allophone of /a/, preceding a cluster beginning with *s*, e.g., *ispu'n-* (M *spwrn-*) “full.” Otherwise it combines with the following *y*, namely *y*, to indicate front vowels, e.g., *ēw* (M *yw*) “one.” Doubling of mater lectiones and *ṭ* (*teth*) like *w*”*xšṭṭ* “words” is not uncommon but is of no phonetic or phonemic relevance (Gershevitch, 1954, §76). In addition to the original 22 letters of the Syriac alphabet, the new characters *f*, *x*, and *ž* were invented. Though *b* (*beth*) and *d* (*daleth*) transcribe /β/ and /δ/, respectively, in most manuscripts a phoneme /γ/ is represented not by *g* (*gimel*) but by ‘(*ain*), which is transliterated as γ. It is only in the materials written in the Syriac script that /δ/ (= *daleth*) and /θ/ (= *tau*) are distinguished, e.g., *θβar-* “to give” (S/M *δβr-*, C *θβr-*) ~ *δūr* “far” (S/M *δwr*, C *dwr*).

Prosodies are only inferred from the rhythmic law: a stress accent falls on the ending of a light stem or on the stem of a heavy stem: *βay-i* “god (nom. sg.)” ~ *βāy-ī* “garden (obl. sg.)” Light stems lose the ending when it become a clitic, e.g., *βay-a* “o lord” ~ *rti-βay* “and o lord.” Several function words including numerals have both independent and clitic forms: *kδ* ~ *kδ'* “if, when,” *δs* ~ *δs'* “ten,” *wy'* ~ *'wy* “(article, loc. sg.),” (*)sty* ~ (*)st* “is.”

Historical phonology. The typologically marked system with voice opposition found only with fricatives is due to the East Iranian sound change in which the voiced plosives and affricate **b*, **d*, **g*, and **ǰ* have become respective fricatives β, δ, γ, and ž even in initial position, and also due to the Sogdian conservatism which preserves the voiceless plosives and affricate **p*, **t*, **k*, and **č* even after a vowel: *βr-* “to bring” < **bara-*, *δ'r* “to hold, have” < **dāra-*, *yr-* “mountain” < **gari-*, C *ž'r* “poison” < **janθra-*; *p* “water” < **āp-*, *w't* “wind” < **wāta-*, *'wt'k* “country” < **awa-tāka-*, *w'c* “to send, release” < **wāča-*. Sogdian conservatism is also seen in its preservation of Old Iranian **θ*, e.g., *xēpθ* (C



xypθ) “(one’s) own” < **xwaipaθya*-. The voiced plosives occur only after nasalized vowels, in which position the Old Iranian voiced and voiceless series have fallen together, as in *saṃg* (S *snk*, cf. M *sngcyk* “stony”) “stone” < **asanga*- vs. *zaṃg* (S *znk*, M *zng*) “kind, sort” < **zanaka*-. Sogdian shared with other East Iranian languages the voicing of the fricatives found in the OIr. clusters *-*ft*- and *-*xt*- to *-*βt*- and *-*γt*-, e.g., (ə)*βta* (C (')*bt*') “seven” < **hafta*, *ōsuytē* (M 'wsuwytyy) “pure” < **awa-suxtaka*-. Another East Iranian feature is the shortening of some long vowels followed by **y* and **w*, e.g., *sayāk* (M *sy*'k) “shadow” < **sāyāka*-, *žu*- (M *jw*-) “to live” < **ǰwa*-.

While consonants are relatively conservative, vowels underwent extreme innovations due to the shift of the stress accent and subsequent loss of unaccented short vowels. The accented syllables *-*aC/āC* were palatalized by the following *-*y* or *-*aya*-, which were later lost, as in *nišēδ* (M *nšyδ*) “to seat, place” < **nišādaya*-, and *neš*- (M *nyš*-) “to be spoiled” < **nasya*-, *frēž* (C *fryž*) “to make straight” < **fra-rāzaya*-. Note also the simultaneous palatalization of *s* and *z* resulting in *š* and *ž*.

Among the combinatory changes of Old Iranian consonant clusters the followings are to be mentioned: *sp* < OIr. **tsw* < IE **k_s*, e.g., 'sp- “horse”; *zβ* < OIr. **dzw* < IE **ǵh_u*, e.g., 'zβ'k “tongue”; *δβ* < OIr. **dw*, e.g., *δβr*- “door” < **dwara*-; *ž* < OIr. **dr*, e.g., C *žwk* “healthy” < **druwaka*-; *š* < **čy*, *sr*, *θr*, e.g., *šw*- “go” < **čyava*-, C *šwn* “hips” < **srauni*-, C *žwšy* “offering” < **zauθraka*- (with assimilation of *z*- to *ž*). Sometimes a metathesis precedes the merger of *δr* and *θr*, e.g., 'rδyšp “banner” < **drafša*-, S *cyrδp'δw* “quadruped” < **čaθru-pāda*-. The dialect variation of *tiray* ~ *siray* “three” (cf. S 'δry, C šy < **θrayah*) found in Yaghnoibi or Modern Sogdian suggests a relatively late development of **θr* to *š*. The development of OIr. **θw* has been extensively discussed by Sims-Williams (2004).

No centralized government having been established in Sogdiana, orthographic rules were not established in Sogdian, and one encounters a considerable number of spelling variations, some of which are due to the historical spelling, yet others to the variations in pronunciation. Most conspicuous is the metathesis of [w, u], which is so common that almost every word containing the sounds has alternative forms (e.g., *δwyth* ~ *δywth* [δuyta ~ δyuta] “daughter”), and this phenomenon is also observed among loanwords (e.g., *smwtr*- ~ *swmtr*- < Skt. *samudra* “world ocean”). When *č* comes into contact with *t*, it often becomes *š*, e.g., *s'ct* ~ *s'st* “it is necessary.” Some forms found in Christian texts in Syriac script seem to reflect vulgar and chronologically later



pronunciations. The loss of nasal vowels before continuants (e.g., C *qθ* [kaθ] < S *knδh* [kaṃθ] “city”) and that of *ə* in the coda position (e.g., C *zyn* [zen] “gold” < S *zyrn* [zeəⁿn]) are peculiar to Christian texts. See also the reduction of the durative particle *skun* to (s)*kən* (C *sqn* ~ *qn* after *-t*), and (s)*k* (C *sq* ~ *q* after *-t*), and that of the future particle *kām* to *kā* (C *q*). Syncope of medial vowels and resultant changes of consonants are also common, e.g., C *wyḍ’rt* < M *wytwδ’rt* “he said” (Sims-Williams, 1985a, p. 97 with n. 21). Total loss of a prothetic vowel preceding consonant clusters are also peculiar to Christian texts, e.g., C *ps’q* ~ S *’ps’k* “garland.” On the spelling variations see also Provasi, 2003.

Morphology. Nouns. In the nominal declension the three numbers and three genders (masculine, feminine, and neuter) of Old Iranian are preserved, although the survival of the neuter is marginal, and many old neuter nouns have shifted to masculine or feminine. Old dual forms have come to be used in the position immediately following a numeral, not only two but also higher numbers. Therefore, this special form is called “numeralive” (Sims-Williams, 1979). Nouns are classified into several declensions. Apart from the distinction between light and heavy stems, a few light stems ending with *-u* (*-w*) inflect differently from ordinary light stems. A considerable number of stems go back to forms extended by the suffixes **-aka* (m.) and **-ākā* (f.) and are conventionally referred to as *aka*-stem and *ākā*-stem. The two stems are often written with historical spellings in Sogdian script, the former with *-(‘)k*, etc. and the latter *-’kh*, e.g., *βamtē* “slave” (S *βnt’k*, M *βndy*), *xānā* “house” (S *x’n’kh*, M *x’n*). Finally there are a small number of indeclinable nouns ending with *-ī*, e.g., *martī* (*mrty*) “man.”

Light stems distinguish six cases: nominative, accusative, genitive-dative, instrumental-ablative, locative, and vocative. Instrumental-ablative forms are never used independently and are always accompanied by the preposition S *cnn/c’wn* (C *cn*) “from” or S *δnn/δ’wn* (C *dn*) “with.” Although the old dative and instrumental merged with the genitive and ablative, respectively, the system is largely in accordance with the pattern of Old Iranian. Most of the plural forms are characterized by the ending *-t*, and these plural stems are treated as feminine singular. Case endings are: masculine nom. *-i*, acc. *-u*, gen.-dat. *-e*, inst.-abl. *-a*, loc. *-ya*, voc. *-a*, numeralive, nom./acc. *-a*; feminine nom./acc. *-a*, gen.-dat./inst.-abl./loc. *-ya*, voc. *-e*. Examples: m. nom. *βay-i* (*βyy*) “god,” acc. *βay-u* (*βyw*), gen.-dat. *βay-e* (*βyy*), inst.-abl. *βay-a* (*βy*), loc. *δβar-ya* (*δβry*) “door,” num. *βay-a* (*βy*), voc. *βay-a* (*βy*); f. nom./acc. *wan-a* (*wn*) “tree,” gen.-dat./inst.-abl./loc. *wan-ya* (*wny*), voc. *δwyte* (*δwyty*, cf. Sims-Williams, 2013);



plural nom./acc. *δβar-ta* (M *δβrṭ*), gen.-dat. /inst.-abl./loc. *δβar-tya* (M *δβrty*), etc. “doors.”

The heavy stem declension does not differentiate masculine from feminine and distinguishes only direct and oblique cases. The oblique case is provided with the ending *-ī* (-y): C *myθ* “day” (sg. dir.), *myθ-y* (sg. obl.); *myθt* (pl. dir.), *myθ-ty* (pl. obl.). The light stem feminine vocative ending *-e* is borrowed by the heavy stem and the pl. stem ending with *-t*, e.g., S *xwtynyh* “o queen,” C *xwt'wty* “o lords” (Sims-Williams, 2013).

In the *aka-* and *ākā-*stems the original intervocalic *-k-* was lost, and resulting hiatus was later contracted (Sims-Williams, 1990). Thus, the m. nom. sg. ending **-aki* first became **-a'i*, and then *-ē*, while the acc. ending *-ō* originates from **-aku* via **-a'u*. In one Christian Sogdian manuscript, C2 (now E27), the original situation after the contraction is well preserved, while in all the other texts, the ending *-ē* is generalized from nom., gen.-dat., and loc. sg. to the other cases. The plural ending *-t* is suffixed to the nominative form, i.e., *-ē-t*, and inflects as a heavy stem: *mrtxmy* “man” (sg. dir.), *mrtxmy-t* (pl. dir.), *mrtxmy-t-y* (pl. obl.). Similarly, in the case of the *ākā-*stem, the effect of vowel contraction leads to a pattern similar to the heavy stem declension: C *šm'r* “thought” (sg. dir.), *šm'ry* (sg. obl.); *šm'ry-t* (pl. dir.), *šm'ry-t-y* (pl. obl.). Tendencies to substitute the nominative for other cases and for the heavy stem to borrow light stem endings are observable. Some animate nouns, mostly light stems, constitute plural forms by means of the ending *-īšt(-ī)*: *βγ-yšt* (dir.), *βγ-yšt-y* (obl.) “gods.” On the origin of the ending, see Sims-Williams (1979). The plural forms of *βr't* “brother” and *δwyt-* “daughter” are peculiar in ending with *-rt*, e.g., *βr't-rt.* and *δwyt-rt.* Archaic pl. gen. ending *-ān* (< OIr. **-ānām*) is sometimes encountered in such stereotyped phrases as *βγ'n βxtm* “most divine of gods (= Skt. *devātideva*).” In the Christian manuscript C5 (now E5), the nominal inflection shows a strong tendency to use the nominative form in all the case functions, and a new oblique case begins to be formed by adding *-ī* to the nom. sg. and similar generalized forms, e.g., nom. m. *rmy* “people” ~ obl. *rmy-y*, nom. f. *wn'* “tree” ~ obl. *wn'-y* (Sims-Williams, 1982). In this stage, the so-called “differential object marking” with this oblique ending is also observed: the new oblique case is employed for marking the direct object which is both definite and human (cf. Yoshida, 2009a, p. 307), e.g., *fšmd'rt qw wyšnt s' xypθ z'ty-y* “He sent (*fšmd'rt*) his son (*xypθ z'ty-y*) to them (*qw wyšnt s*).”

The inflections of Sogdian nouns and adjectives are almost identical, and what is described about nouns applies also to adjectives, which, however, lack



numeral forms. The productive suffix of the comparative is *-(y)str*, e.g., *yw'ncyk-str* “more sinful.” Somewhat obsolete is the suffix *-tr*, which is not attached to stems derived with suffixes or compounded forms: *δwr-tr* “farther,” but M *mndyrβ'k-str* “more stupid.” The adjective and adverb possess a special formation called “elative” with the meaning “so much ~, very ~.” It is formed by means of the prefixes *c'*– “how much” *w'*– (*wt*– before *s*–) “so much,” and suffixes *-t*, *-(')st*, in various combinations: *w'-z'ry* “so miserable,” *c'-z'ry-st* “id.,” etc. The old superlative ending *-tm* becomes obsolete and survives only in such fixed expressions as *βγ'n βxtm* and forms extended by another suffix, cf. *'sk'tmcyk* “highest, most.”

Word-formation of nominal stems. Productive ways of forming new words are derivation by affixes and compounding. Derivational prefixes are not so common as suffixes. Many suffixes are *aka*-stems, of which the feminine counterpart ends not with *-k'h* but *-c*. The most productive suffixes are the following: *-yny* (f. *-ync*) “(made out) of ~,” e.g., *δ'rwkync* < *δ'rwk* “wood”; *-cyk*, e.g., *yrcyk* “of mountain” < *yr*–; *-mync*, e.g., *'yncmync* “female” < *'ync* “woman.” Nouns provided with the suffix *-nc* denote the female counterparts of titles and occupations, e.g., *nywš'k'nch* “auditorice” < *nywš'k*. For its origin, see Avestan f. *ahurānī*– < m. *ahura*– “lord.” Abstract nouns are derived with such suffixes as *-wny* (*-ōnī*), *-y'k* (*-y'* with heavy stems), and *-wy* (*-āwē*), e.g., *t'ywny* “theft” < *t'y* “thief,” M *rwxs'ny'k* “lightness, light,” S *ptpt'yn'w'k* “isolation” < *ptptyn* (Sims-Williams, 1981b). The suffix *-mnty* (*-āmamtē*), which derives verbal nouns from the present stem, is very common in Manichean and Christian texts, but rare in Buddhist texts, e.g., C *'ys'mnty* < *'ys* “to come.” Productive prefixes are *'w*– “co-,” and privatives *n'*–, *(')pw*–, and *mnt*–, e.g., *'wn'm* “namesake” < *n'm* “name,” M *n'-pδkcyq* “unlawful,” M *mndyrβ'k* “foolish,” *'pw-ptšm'r* “innumerable.” The most common compounds are *bahuvrīhis* and agent nouns/adjectives consisting of the present stem as the second member; both types end with the *aka*-suffix, e.g., M *xii-ryty* (*δwāts-rītē*, cf. *ryt* “face”) “having twelve faces” and S *yp'k-βr'k* “furious” (*yp'k* “anger” + *βr*– “bear”). On Sogdian compounds in general, see Gershevitch (1945).

Numerals. Cardinal numbers except for “2” are indeclinable. Oblique forms ending with *-nw* are occasionally encountered: 1 *'yw* [ēw] (M *'yw*, C *yw* [yō]), 2 m. *(')δw()*, f./n. S *'δwy*, C *dwy*, obl. *δyβnw*, 3 M *'δryy* (C *šy*), 4 M *ctf'r*, 5 *pnc*, 6 *wxwšw* ~ *wxwšwnw* (C *xwšw*), 7 S *'βt()* ~ S *'βtnw*, 8 S *'št()*, 9 S *nw()*, 10 S *δs()*, 20 C *wyst*, 30 *šys*, 100 C *stw*, 200 C *dwyst*, 1000 *z'r*, S *1-LPw-nw* ([zārnu]), 10,000 *βrywr*. The light stem forms without *-'* like *βt* are originally proclitic forms. In



the cardinal numbers “11” to “19,” one digit figure is followed by δs “10,” which itself is reduced to [ts]; cf. B *pnc* δs , M *pnc* δsyh (obl.), C *pncc* “15.” Except for 1st (M *’ftm*-, etc.), 2nd ($\delta\beta ty$ -, etc.), and 3rd ($\delta’sty$ -, etc.), ordinals are formed with the suffix *-m* (LS *-my*) or *-myk*, e.g., S $\beta tmy \sim \beta tmyk$ “7th.”

Pronouns. Personal pronouns and their enclitic forms are as follows: 1st sg. dir. (δ)*zw*, obl. *mn’*, encl. *-my*; 2nd sg. dir. *tyw*, obl. *tw’*, encl. *-fy* (in some texts acc. *-f*, gen.-dat. *-t(y)*); 1st pl. *m’x*, encl. *-mn*; 2nd pl. *šm’x*, encl. *-fn*; 3rd sg. encl. *-šy/šw*, pl. *-šn*. In some texts M/C *t’m’* (S *t’m’kh*) and M/C *t’f* (S *t’β’kh*) function as direct object forms of the 1st and 2nd person sg. pronouns. In the latest stage of the language attested in the Christian text C5 (now E5), the secondary oblique ending *-y* is also attached to the personal pronouns: *mn’y*, *šm’xy*, *m’xy*, *wyny* (Sims-Williams, 1989a, p. 186). Enclitic forms are attached to the first element of the sentence irrespective of their function. On the independent forms of the 3rd person pronoun, see below.

Sogdian demonstratives are characterized by the triple system of deixis depending on the three persons, each showing two bases: *y-/m-* “this (with me),” *š-/t-* “that (with you),” and *x-/w-* “that (with him, etc.)” (Sims-Williams, 1994). The three are also seen in such adverbs as *mδy* “here (by me),” *tδy* “there (by you),” and *wδy* “there (by him, etc.).” *y-/m-* and *x-/w-* bases have both anaphoric and deictic uses, but *š-/t-*, of which the instances are not many, seems to be employed only as deictic. Among the demonstratives, simple and extended forms are distinguished; the former function as the article and 3rd person pronouns, while the latter are extended from the former and serve as strong demonstratives. The masculine (and feminine) singular forms of *x-/w-* deixis are as follows: nom. *xw* (f. *x’*); acc. *’w* (f. *w’*); gen.-dat. *wyny* $\sim$ *’wyn* (*wy’* $\sim$ *’wy*); loc. m./f. *wy’* $\sim$ *’wy*. The corresponding ideograms are based on ZK, e.g., ZK = *xw*, ZK*wy* = *’wy*. In late texts articles undergo phonetic reduction and are prefixed to the following nouns: *w-m’n* “mind” (acc.), *n-yr’m* “wealth” (gen.-dat.), *y-m’ny* (loc.). The extended forms are nom. *xwny* or *xwnx* (f. *x’n*); acc. *’wnw* (f. *w’n*); gen.-dat. (*w*)*nyw’nt*; loc. *wy’wnt*. Another extended form is provided with the element *-ēδ*, e.g., *x-yδ* and *w-yδ*. In the texts written in Sogdian script (δ)*xw* = ZK is sometimes used as a copula. The corresponding *y-/m-* forms are nom. *yw* = ZNH, *ywny* (f. *y’n*); acc. *’mw*, *mwnw* (f. *m’n*); gen.-dat. *’myn*, *nym’nt*; loc. *my’*, *my’mnt*; *yδ* and *myδ*. It is to be noted that the employment of the article is declining and one finds many fewer articles in Christian texts than in the Mug documents of the early 8th century and the contemporary Buddhist texts. The Sogdian prepositions are portmanteau



forms containing a weak pronoun or article (cf. below) and they are scarcely followed by another article. This indicates that no distinction is made between definite and indefinite by means of the articles, of which the function is more syntactic than semantic/pragmatic. On the Sogdian weak demonstratives or articles, see Wendtland (2011). An adjective *wysp-* “all” sometimes takes pronominal endings, e.g., gen.-dat. *wyspny*, ins.-abl. *wyspn’δ*, *wyspn’c*, and pl. nom. *wyspy*, gen.-dat. *wyspyšnw*.

Sogdian pronouns also have portmanteau forms containing a prepositional element: with 1st and 2nd sg. pronouns: S *c-’m’kh* “from me,” *c-’β’kh*, *δ-’m’kh*, etc. (the *t-* element of *t’m’*, etc. mentioned above is an obsolete preposition *t’()* “to”); with articles: *c-’wn* (< OIr. **hača awanā*; Sims-Williams, 1990, p. 277, n. 5) “from the ~,” *δ-’wn* “with the ~,” *pr-w* “on the ~,” etc.; with strong demonstratives: *pr-yw’nt/pr-ywyδ* “on that/those,” etc.

No special reflexive pronoun is known, but a feminine noun *γryw* “body” functions as a kind of reflexive. The forms derived from OIr. **xwa-* are *xwty* “self” and *xypδ* (C *xypθ*); the former emphasizes a personal pronoun (often not expressed), e.g., *’zw xwty* “I myself,” while the latter expresses possession “(one’s) own.”

Interrogatives function also as relatives, which are often provided with the conjunction (*’t*(y), on which see below: *ky* (obl. *ky’*, abl. *ckn’c*) “who,” *cw* “what,” *kδ* “when,” *kw* “where,” *c’nw* “how,” *c’f* “how much,” etc.

Indefinite pronouns are *’δy*, *’yδy* “someone, anyone,” *nyδy* “nobody,” *’δc*, *’yδc* “something, anything,” *nyδc* “nothing.” Reciprocals are expressed by the combination of *yw* “one” and an inflected form of *δβty-* “the second.”

Verbs. Each Sogdian verb has two stems: present and past. Historically the former originates from the Old Iranian present stem and the latter, which always ends with *-t*, from the past participle derived with the suffix **-ta*. Therefore, the two stems may sometimes differ remarkably: e.g., pres. stem (*k*)*wn-* “to do, make” ~ past stem S *’krt-*, (C *qt-*) going back to OIr. **kṛnau-* and **kṛta-*, respectively. Several common verbs derive their present and past stems from different roots, e.g., *w’β/wyt-* “to say,” *’βr/’yt* “to bring.” The productive way to derive past stems from the present is to add a suffix *-(’t)*. Some irregular verbs have both forms, e.g., S *’prs-* (C *ps-*) “to ask” and its past stems M *fšt-* and S *ps’t*.



Pairs of transitive/causative and intransitive/passive stems are frequently met with, e.g., *xwyr* “to feed” ~ *xwr-* “to eat,” *βr-* “to bring” ~ *βyr-* “to be brought.” Some of the intransitive/passive forms derive from the old inchoative stems (Weber, 1970), e.g., *swc/swyt-* “to burn” ~ *swxs-* “to be burnt,” *wyryš/wyr’t* “to arouse, wake” ~ *wyr’s* “to wake up.”

In addition to the present and past stems, certain verbs possess distinct imperfect stems characterized by the preservation, or analogical extension, of the Old Iranian augment **a-*. In principle, only those stems containing preverbs undergo certain processes by which the old augment (inserted between a preverb and a stem) and a final vowel (**-i* or **-a*) of the preverb are fused into *-ī-* or *-ā-*: e.g., *patyōš* (*ptywš*) “to hear” > *patīyōš* (*ptywš*), cf. OIr. **pati-a-gauša-*; *framāy* (*frm’y*) “to order” > *frāmāy* (*fr’m’y*), cf. OIr. **fra-a-māya-*. Contrary to expectation, verbs containing a preverb *β-* (< OIr. **abi*) show the augment *-ā-* rather than *-ī-*, e.g., *βž’y-* “to increase (intransitive)” > *β’žy* (< OIr. **abi-jawya-*); similarly, those beginning with **us/uz-* take *-ī-*, e.g., *()zw’rt* “to turn” > *zywrt* (< OIr. **uz-warta-*). Present stems beginning with a preverb *n-* (< OIr. **ham-*) form their imperfect stem by adding *m-* to the present, e.g., *’nxz* “to rise” > *mnxz*. This *m-* was later extended to verbs beginning with the preverb *ā-* (< OIr. **ā-*), e.g., *βr* “to bring” > *m’βr*. Analogical extension to etymologically simple verbs is also known: *sn’y* “to wash” > *syn’y* (< OIr. **snāya-*).

Forms based on the present stem. Sogdian has six moods: indicative, subjunctive, optative, injunctive, irrealis, and imperative. The so-called precative has been shown to be the middle conjugation of the optative (Sims-Williams, 1989a, p. 188; Yoshida, 2009a and 2009c). The tense and aspect system comprises present, imperfect, optative imperfect, preterit based on the past stem, *’z*-imperfect, and *yq* (*w*)*m’t* imperfect, the latter two being productive only in some Christian Sogdian texts. The perfect tense and the periphrastic passive are formed from the past participle (i.e., the past stem extended by the *aka*-suffix) and auxiliary verbs. Old middle endings are almost all displaced by the active, only a few of them having survived, among which is the occasionally encountered *-ymn* (1st pl.). However, 2nd sg. and 3rd sg. optative middle and 3rd sg. imperfect middle endings are formally so salient that they serve as starting-points for the analogical development of new paradigms. The old force of the middle voice is only perceivable in the 3rd sg. present with the ending *-ty*, which conveys passive meaning when used with transitive verbs, e.g., *wyn-ty* “(he) is seen” as against active *wyn-t* “(he) sees.” Durative and future meanings are conveyed, respectively, by particles



-(*skwn*) (in a few texts occasionally -*štn* as well; cf. Benveniste, 1966, who compares it with Yaghn. -*išt*) and -*k'm*, e.g., *šw'm-skwn*, *šw'm-štn* (cf. Yaghn. *šawomišt*) “I am going” and *šw'm-k'm* “I will go.” While *k'm* is attested in the Ancient Letters, *skwn* or *štn* does not occur there.

Present indicative forms (for the sake of convenience, idealized forms based on sample stems *kun-* “to do, make” and *patyōš* “to hear” are given): 1st sg. *kun-ām* (*patyōš-a/ām*), 2nd sg. *kun-e* (*patyōš-e*), 3rd sg. *kun-ti* (*patyōš-t*), 1st pl. *kun-ēm* (*patyōš-ēm*), 2nd pl. *kun-ta/θa* (*patyōš-θ/ta*), 3rd pl. *kun-aṃt* (*patyōš-aṃt*); imperative: 2nd sg. *kun-a* (*patyōš-ø*), 2nd pl. *kun-θa/ta* (*patyōš-θ/ta*); imperfect: 1st sg. *kun-u* (*patyōš-u*), 2nd sg. *kun-i* (*patyōš-i*), 3rd sg. *kun-a* (*patyōš-ø*), 1st pl. *kun-ēm* (*patyōš-ēm*), 2nd pl. *kun-θa/ta* (*patyōš-θ/ta*), 3rd pl. *kun-aṃt* (*patyōš-aṃt*); optative: 1st sg. *kun-ē*, 2nd sg. *kun-ē/ya*, 3rd sg. *kun-ē*, 1st pl. *kun-ēm*, 2nd pl. *kun-ēθ*, 3rd pl. *kun-ēmṭ*; subj. 1st sg. *kun-ān* (*patyōš-an/ān*), 2nd sg. *kun-a* (*patyōš-a*), 3rd sg. *kun-āt* (*patyōš-at/āt*); injunctive: 1st sg. *kun-u* (*patyōš-u*), 2nd sg. *kun-i* (*patyōš-ø*), 3rd sg. *kun-a* (*patyōš-ø*), 1st pl. (not attested), 2 pl. *kun-θa/ta* (*patyōš-θ/ta*), 3 pl. *kun-aṃt* (*patyōš-aṃt*). Optative middle endings are: 1st sg. -*ēm/ētu*, 2nd sg. -*ēš/-ēta*, 3rd sg. -*ēt/-ētē*, 1st pl. -*ētēman*, 2nd pl. *ēšθ(a)*, 3rd pl. -*ētēmṭ*; irrealis endings are: 1st sg. -*ōtu*, 2nd sg. -*ōta*, 3rd sg. -*ōtē*, 2nd pl. -*ōtēšta*. In the *āz*-imperfect and imperfect middle, the imperfect endings follow the present stem enlarged respectively by -*z* and -*t*: *δār-āz-u* “I was holding,” *ās-t-u* “I took.” However, only two verbs, (*k*)*wn-* and *s* “to take” show the imperfect middle (Sims-Williams, 1996b, p. 177, n. 11). Optative forms sometimes denote durative/iterative past and are referred to as “imperfect optative,” which often takes the imperfect stem rather than the present: *w'pt'y* “was falling, fell repeatedly” < *wpt*.

Nonfinite forms based on the present stem are present participles formed with -(*ē*)*nē* (f. -(*ē*)*nč*), e.g., *S pt'wn'k* “enduring,” *S δbr'yn'k* “giving,” future passive participles ending with -(*i*)*čik* or -*i*, e.g., *M swmbcyk* “to be bored,” *S βyry* “to be found,” and gerunds formed by the suffix -*kya/kī*, e.g., *βrky'* “having brought,” *w'βky* “having said.” Infinitives derived from the present stem are also found. Their formation and usage vary from text to text (Yoshida, 1979). One type of them, the so-called *pr*-infinitive consisting of the preposition *pr* and the present stem, is common only in Christian texts. In some Christian Sogdian texts, the present participle ending in -*ēk* or -*esk* construes with the auxiliary verb (*w*)*m't*, which itself is the preterit form of the copula, and the combination denotes the durative past, namely *yq* (*w*)*m't*-imperfect, e.g., *C šwyqm't* “he was going.” In the texts in which *āz*-imperfect is attested the *yq*



(w)m't-imperfect is virtually absent (Yoshida, 1980). The Old Iranian present participle ending with *-ant is no longer productive. The form extended with the *aka*-suffix is found in some fossilized adjectives, e.g., *žuwaṃtē* (M *jw'ndy*) “living.”

Forms based on the past stem. The intransitive or passive preterit consists of the past stem and the enclitic forms of the copula, except for the 3rd sg., which is nothing but the nom. sg. form of the past stem (idealized forms based on S *βw-/krt*- “to become” and *ʾys/ʾyt* “to come”): sg. 1st *əkt-im*, 2nd *əkt-iš*, 3rd *əkt-i*, (both masculine and feminine, but occasionally f. sg. *əkt-a*; *āyat-ø*), pl. 1st *əkt-ēm*, 2nd *əkt-(aṃ)sθ*, 3rd *əkt-aṃt*. The corresponding transitive preterit is formed with the auxiliary verb *δ'r* “to have.” The light stem takes the old acc. sg. ending -u while the heavy stem has no ending: 1st sg. *žaytuδāram* (M *jyṭw-δ'rm*) “I held (M *δ'r/jyṭ*-),” *patxrūt-δāram* (M *ptxryṭδ'rm*) “I hired (*ptxryn/ptxryṭ*),” etc. The modal forms of the preterit tense are formed by inflecting the auxiliary verb, e.g., S *wm't-t* “(he) would have been (subj.),” M *wjyystδ'rn* “(if) I should have settled down (subj.; M *wjyynd/wjyyst*).” Meager survivals of the so-called “ergative construction” based on the past stem of the transitive verbs are also encountered: *ʾHRZYm ZK dykh wyth* “I (-m) saw the letter (*dykh*) (*wyn/wyt* “to see”; cited from one of the Ancient Letters).” The potentialis is a periphrasis consisting of a past stem in -t (heavy) or -tu/ta (light) plus auxiliary verb (*k*)wn- “to make” (trans.) or *β(w)*- (intrans./pass.). On the origin of the construction, see Sims-Williams (2007b). Two senses can be distinguished, one “potential” (e.g., M *ṭwyṭ' kwnyy* (3rd sg. opt.) “he might be able to pay”), the other “anterior” (e.g., S *c'n'kw ... wytw wnt'* “when he had said”). Anterior meaning is also expressed by the non-finite construction based on the potentialis: *čan ~tu/ta kārī*, e.g., S *mn' cnn nyrβ'n wyt'rt k'ry* “after my passing into nirvāṇa (*wytr-/wytrt* “to depart”),” S *cnn pwyt'-k'ry* “after cooking (it) (*pc-/pwyt*- “to cook”).”

The past participle is derived from the past stem with the *aka*-suffix, e.g., m. *parθaytē* ~f. *parθay(t)č* < S *prδ'ync/prδ'yṭ*- “to pull.” The periphrastic passive and intransitive perfect are formed with the past participle (agreeing with the subject) and the auxiliary verbs, which are *β(w)*- (passive) and existential verbs (perfect), e.g., M *sfrytyṭ wβ'nd* “they are created (pass.)” and S *tyṭ'y 'skw'y* “you have entered (perfect).” The active perfect with the auxiliary verb *δ'r* is attested but not common: S *ptyrβ't'k δ'r'nt* “they have received.” More often the passive construction is employed: S *δβ'rt'y 'st't* “(if) it should have been given.” The past infinitive ends with -te (light) or -t (heavy) and always follows the



main verb: e.g., S *s'ct* 'PZYšy ... *cšmt'* *knt* "it is necessary (*s'ct*) to dig out (*kn-/knt*) his eyes." On the Sogdian infinitives, see Yoshida (1979).

Indeclinables. Both prepositions and postpositions are common. Those which are construed with accusative case of the light stem are *pr*, *prw*, *prm* "in, on, by, for," and *()kw* "to, toward," which has replaced the obsolete *t()*. *cnn*, *c'wn* "from" and *δnn*, *δ'wn* "with" take the instrumental-ablative. Somewhat rare is *wsn* "for the sake of" (Sims-Williams, 2001b).

The most common postpositions are *s'r* (C *s'*) "toward, from," *pr'yw* (C *prw*; < *pr* + *yw* "lit. in one") "together with," *prm* "up to, until," and *pyδ'r* "for the sake of, because of" which are often preceded by *()kw* or *cnn/c'wn*, e.g., *kw δynh s'r* "to the church," *cnn rym'yš pyδ'r* "because of rebuke." Many adverbs and oblique case forms of nouns function as postpositions, e.g., *myδ'ny* "among" < *myδ'n* "middle, waist," *ryty* "in front of" < *ryt* "face," *cyntr* "inside," *cwpr* "upon." Thus, Sogdian postpositions constitute an open class in contrast to the prepositions.

Sentences are negated by either *nē* (M *ny(y)*, C *ny*) or *nā* (M/C *n'*) preceding finite verbs, both being masked by the ideogram *L'*. In the case of the perfect and periphrastic passive, *nē* precedes the auxiliary verb. While *nē* negates proposition, *nā* is a prohibitive particle: M *nyy w'β'mk'm* "I shall not speak," M *kpyy nyy ny'tδ'rt* "he did not catch the fish," C *yty ny bwtq'* "it will not be taken," C *n' ps'* "don't ask!" The negation of the imperfect is formed not by negating the imperfect verb, but by negating the present indicative or injunctive form, occasionally preceded by an enclitic element S *-β(y)* (Sims-Williams, 1996b): S *rty-βy L' δwry z'yh šwt* (pres.) "He did not go far."

Conjunctions. Coordinate conjunctions are *()rt(y)* (= 'HRZY) which marks the beginning of the clauses, *t(y)* (= 'PZY, ZY) "and," and *kt'r* (= 'WZY) "or." Subordinate conjunctions are: (a) preceding the main clauses: *cw* "if," *c'nw* "when, while, since," *kδ()* "if," *mnt* "when, while," *kw prm*, *kwδprm* "as long as"; (b) following the main clauses: *t(y)* = ZY "that," *c'nw ()ty* "as (= like), than," *p'r(w)ty* "for, (not...) but," *pr'w t(y)* "because," and *yw'r (ty)* "however." On the element *ty*, etc. (= ZY), see below. Combinations of *kδ* and *t(y)/-wty* and *m'δ* "thus" and *t(y)* come to be new subordinate conjunctions *kt* (C *qt*) "that" and *m't* "that" respectively. Somewhat rare is *twty* "that, then, and" (combination of *t(y)* "and" and the complementizer, on which see below).

Syntax and phraseology. A syntactic feature peculiar to Sogdian (and for that



matter Bactrian as well) is that in each clause (both main and subordinate) an enclitic complementizer (')*t(y)*, -*wty*, *yty* (= 'PZY, ZY) stands in the second position, to which other enclitic elements are added (Sims-Williams, 1985b). On the usage of the corresponding ideogram ZY, see Yakubovitch (2005). For example, the most frequent conjunction *rt(y)* = 'HRZY marking the beginning of a clause consists of an obsolete adverb *r-* (cf. Khotanese *rro* “also”) and -*ty*. This conjunction is also affixed to the first element of a direct quotation, e.g., S KZNH *pt'yškwy* <wγš' ZY βγ' xwt'w p'rZY γr'nh 'krt'ym> “She said thus (to him), ‘Rejoice, o lord king, because I have become pregnant’” (Weber, 1971). Similarly, the above mentioned subordinate conjunctions (e.g., *p'r-(w)ty*, etc.) contain the complementizer. This is also true with the relative clauses, where the relative pronouns or adverbs are often followed by it: M *xii δβrt' ky-'ty wy' sm'nyty 'skwnd* “Twelve doors which exist on the heavens.” During the period when Sogdian was attested use of the complementizer declined and in later texts it only survives in fossilized forms like C *qyt* (< *ky 'ty*), M/C *kt/qt* (< *kδ-wty*), etc. Similarly, less and less examples of *rt(y)* are attested in later texts.

Word order. Sogdian is basically an OV language where heads follow the dependent elements. The basic structure of the Sogdian sentences is SOV: M *r'ty xww mry'rty xypδ'wnd m'yδ pwskf'ty ww 100 δyn'r zyryn t'wj* “Thus under such constraint the owner (*xypδ'wnd*) of the pearls (*mry'rty*) paid (*t'wj*) the hundred gold dinars (*δyn'r zyryn*).” However, Sogdian is far from being a consistent OV language and attests a number of counterexamples. For example, both prepositions and postpositions are common and the relative clause always follows the head noun: M *wyspw 'rk cw 'ty-my t'γw frm'y'y* “All the work (*wyspw 'rk*) which (*cw 'ty*) you order me.”

As a Middle Iranian language, Sogdian still observes “Wackenagel’s law,” by which enclitics occupy the second position in the sentence. In classical Sogdian the second position is usually occupied by the complementizer (')*t(y)*, and the enclitics are added to it: M *p'r-ty-šy xw wyn' jnyy frm't-δ'ryy* “But you ordered (*frm't-δ'ryy*) him (-*šy*) to play (pres. inf. *jnyy*) the lute (*wyn'*).” Among the enclitics are the above-mentioned -*β(y)*, a hypothetical particle -*n* (Sims-Williams, 2007a, p. 192b), and other clitic elements. These enclitics can pile up: 'HRZYmβc *δykh* ”γt “A letter (*δykh*) came to me (-*m-*) from you (-*β-c*).” After this rule of syntax became no longer obligatory, enclitic pronouns became more independent and stand without their host: S ... *šw kδ'c L' w'c'yδ k'm šw ms L' ptxwyδ* “Never let him escape! Moreover, do not kill him!”

With regard to the relaxed agreement phenomena in Sogdian, cf. Sims-



Williams's (1989, p. 190) remark: "The rules of agreement which apply to light stems are relaxed in the case of heavy and contracted stems, leading to phenomena such as group inflexion. To a large extent the suffixes *-ī* (obl.) and *-t* (plural) are treated as optional, being often omitted where clarity is unimpaired." "Group inflexion" is a phenomenon where only the last in a series of (usually asyndetically coordinated) words is inflected (Gershevitch, 1954, §§1639-43): e.g., M. *cn 'nrx pxrty βyq* "with the exception of fixed stars (instead of pl. obl. form *'nrxty*) and planets." The rule of group inflexion is not compulsory either.

Honorific and polite expressions are collected by Yoshida (2006). Among them are *frm'y* "talk" (the subject being someone higher in rank), lit. "order," and its opposite *ptškw'y*, lit. "entreat"; both mean "to say, speak." The addressor's humble feeling is also expressed by the 1st sg. injunctive and a verb *rxn-* "dare": *S rtyñ 'zw c'ñw rxn'w 'PZYñ ywn'k 'šβr'ck' 'prs'w* "How dare I ask (the Buddha) about this matter?" Note also the use of the hypothetical particle *-n* in this sentence.

Language contact and lexicon. The Sogdian lexicon consists of three groups of words: (a) those inherited from Old Iranian, (b) loanwords, and (c) foreign elements or nonce forms temporarily appearing in texts. Some of the native Iranian elements show distinctively East Iranian features, e.g., S *knδh* "town," *kp-* "fish," and C *myθ* (S, M *myδ*) "day." Fully assimilated loanwords mainly originate from Western Iranian (Middle Persian and Parthian) and Indian. For example, the names of the seven days of a week are of Middle Persian origin: *myr* "Sunday," *m'x* "Monday" (see [HAFTA](#)). Other examples include *š'nš'y* "king of kings" and *sp'ntn* "grain of mustard" (cf. original Sogdian form *šywšpδn*). In view of the pronunciation *-c*, *rwc* "day (of a month)" was borrowed from Old Persian *raučah* rather than Middle Persian *rōz*. Parthian forms are more common: *swkβ'r* "monk," *msyδr* "presbyter," *βypwryc* "divine maiden," etc. The word *myδβ-* "minister" (< Parth. *mayβeδ*) sometimes construes with yet another Parthian loan word: *wz'rkt myδβt'* "great ministers," cf. Parth. *wzrg* (Sims-Williams, 1983b, p. 44). Of Indian origin are *mwδy* "price," *rtn-* "jewel," *pr'ny* "insect," *swmtr-* "ocean," *mkr-* "monkey," *škr-* "sugar," *prm'n* "reliable," etc. (Sims-Williams, 1983a; idem, 2007, p. 252). Two doublets of Sogdian and Bactrian words are found: Sogd. *cxr-* ~ Bactr. *sxr-* "wheel," Sogd. *rδδ-* ~ Bactr. *rx'kh* "cart." An Iranian-Indian hybrid word *s'rtp'w* "caravan leader" (cf. Skt. *sāṛthavāha-* "id.") may also originate from Bactrian (Sims-Williams, 1996a, p. 51, n. 37). Greek loanwords are not few, but their immediate origins are



obscure: *δyδym* “diadem,” *nwm* “law,” M *qpyδ* “shop” (Sims-Williams, 1996a, p. 51, n. 39); *δrxm*– “drachma” is from Greek, but *δyn’r* “dinar” from Latin. Foreign elements are Sanskrit forms in Buddhist texts (Provasi, 2013), Western Middle Iranian lexical items in Manichean texts, and Syriac words in Christian texts. (On the Syriac elements, see Sims-Williams, 1988.) The foreign elements are quite numerous in translations, and their number seems to depend on the scholarship of each translator. Chinese elements are found, but they seem to be cultural words shared by other Central Asian languages: *šnk* “pint” (Chin. *sheng* 升, Uighur *šing*, Khotanese *šim̐ga/šam̐ga*, Tocharian *šaṅk*), *tym* “inn” (Chin. *dian* 店, Uighur *tem*, Pers. *tim*), *mk’* “ink” (Chin. *mo* 墨, Uighur *mäkkä*) (Yoshida 1994, p. 379; Sims-Williams, 1996a, p. 62).

Sogdian lent numerous words to Old Turkic, in particular Uighur: *ažun* (< *žwn*) “living being,” *tamu* (< *tm-*) “hell,” *užik* (< M *’wjk’k*, *’wj’k*) “letter, character,” etc. Sogdians’ cultural influence may be seen in the word for “paper” *k’yδ(y)*, which was borrowed into several languages: Uighur *kägdä*, Persian *kāgād*, Central Asian Sanskrit *kākali*. However, this word itself seems to have been borrowed from Greek *chartés* (Sims-Williams, 1996a, p. 62). Since the New Persian literary language came into being among the speakers of [Eastern Iranian languages](#), including Sogdian, a number of Sogdian words found their way into New Persian: *čayz* (< *cyz*) “frog,” *nayz* “good, excellent” (< *nyz-*), *pasāk* (< *()ps’k*) “garland,” *setāγ* (< *st’γ*) “childless.” It is curious to see Sogdian *θ* and *δ* merge into *l* in New Persian loanwords, e.g., *lenj* “to pull out” < **θēnč*, cf. *zβāk-θēmčē* (S *zβ’k-δync’k*) “one who pulls out a tongue,” *lastpardarak* “handkerchief” < C. *dstprtry* “id.,” see [SOGDIAN LANGUAGE ii. Loanwords in Persian](#)).

Varieties of Sogdian and Yaghnobi. The Sogdian language documented in the bulk of materials handed down to us most probably represents the standard variety spoken in the area surrounding Samarkand during the 7th to 10th centuries; the Bukharan dialect cited by Islamic writers is slightly different (Henning, 1958, pp. 85-86; Sims-Williams, 1989b, pp. 165-66). Linguistic features of the much earlier language are attested in the Ancient Letters of the early 4th century (see [Ancient Letters](#)). The so-called “Turco-Sogdian” is one of the latest varieties of the language, which shows very strong Turkish influence, not only in vocabulary but also in syntax. It was spoken (or merely written?) by those who were bilingual in Sogdian and Old Turkic (Sims-Williams and Hamilton, 1990, pp. 10-11; Sims-Williams, 2008; Yoshida, 2009d, 2011a; Sundermann, 2016). The wide range of linguistic differences that once



existed in Sogdiana may be inferred by comparing Sogdian with Yaghnobi, the so-called “Modern Sogdian.” For example, the formation of imperfect stem by adding the augment *a-* to any present stem (Yaghn. *piraxs-* > *a-piraxs-* “left” vs. Sogd. *prxs-* > *p’rxs* “id.”) and the 3rd pl. ending *-or* (*a-wen-or* “they saw” vs. *wyn-’nt*) cannot be attributed to the linguistic change from Sogdian to Yaghnobi. In the formation of the imperfect stem, Choresmian is much closer to Sogdian than to Yaghnobi (MacKenzie, 1974). On the other hand, it may be worth noting that Yaghnobi shares the 3rd pl. *-r* ending with the neighboring Choresmian and Khotanese.

BIBLIOGRAPHY

Émile Benveniste, *Essai de grammaire sogdienne. Deuxième partie: Morphologie, syntaxe et glossaire*, Paris, 1929.

Idem, “Le verbe *stā-* comme auxiliaire en iranien,” *Acta Orientalia* 30, 1966, pp. 45-49.

Mikhail N. Bogolyubov and Ol’ga I. Smirnova, *Sogdiiskie dokument s gory Mug*, III, Moscow, 1963.

Mary Boyce, *A Catalogue of the Iranian Manuscripts in Manichaean Script in the German Turfan Collection*, Berlin, 1960.

Desmond Durkin-Meisterernst and Enric Morano, *Mani’s Psalms. Middle Persian, Parthian and Sogdian Texts in the Turfan Collection*, Turnhout, 2010.

Robert Gauthiot, *Essai de grammaire sogdienne. Première partie: Phonétique*, Paris, 1914-23.

Ilya Gershevitch, “Sogdian Compounds,” *Transactions of the Philological Society*, 1945, pp. 137-49.

Idem, *A Grammar of Manichean Sogdian*, Oxford, 1954.

Badri Gharib, *Sogdian Dictionary*, Tehran, 1995.



Frantz Grenet and Étienne de la Vaissière, “The Last Days of Panjikent,” *Silk Road Art and Archaeology* 8, 2002, pp. 155-96.

Walter B. Henning, *Ein manichäisches Bet- und Beichtbuch*, Abhandlungen der Preußischen Akademie der Wissenschaften, Phil.-hist. Klasse, 1936 [1937].

Idem, “Sogdian Loan-Words in New Persian,” *BSOS* 10, 1939, pp. 93-106.

Idem, “Mitteliransich,” in B. Spuler, ed., *Hanbuch der Orientalistik*, I/4, *Iranistik I: Linguistik*, Leiden, 1958, pp. 20-130.

György Kara, “Aramaic Scripts for Altaic Languages,” Peter T. Daniels and William Bright, eds., *The World’s Writing Systems*, New York, 1996, pp. 536-58.

Sergej Klyashtorny and Vladimir A. Livshits, “Une inscription inédite turque et sogdienne: la stèle de Sevrey (Gobi méridional),” *JA* 259, 1971, pp. 11-20.

Idem, “The Sogdian Inscription of Bugut Revisited,” *AOH* 26, 1972, pp. 69-102.

Kogi Kudara, Werner Sundermann, and Yutaka Yoshida, *Irango Danpen Shūsei/Iranian Fragments from the Otani Collection*, 2 vols., Kyoto, 1997.

Étienne de la Vaissière, *Sogdian Traders. A History*, tr. James Ward, Leiden, 2005.

Vladimir A. Livshits, “A Sogdian Alphabet from Panjikant,” in Mary Boyce and Ilya Gershevitch, eds., *Henning Memorial Volume*, London, 1970, pp. 256-63.

Idem, *Sogdiskaya èpigrafiika Sredneĭ Azii i Semirech’ya*, Saint Petersburg, 2008; English tr. as *Sogdian Epigraphy of Central Asia and Semirech’e*, *Corpus Inscriptionum Iranicarum Part II, Vol. III*, London, 2015.

Pavel Lurje, *Personal Names in Sogdian Texts*, Vienna, 2010.

David N. MacKenzie, “Khwarezmian Imperfect Stem,” in *Mélanges linguistiques offerts à Émile Benveniste*, Paris, pp. 389-95.

Idem, *The Buddhist Sogdian Texts of the British Library*, (*Acta Iranica* 10), Tehran/Liège, 1976.

Dieter Maue and Nicholas Sims-Williams, “Ein sanskrit-sogdische Bilingue in Brāhmī,” *BSOAS* 54, 1991, pp. 486-95.



Moqaddasi, *Aḥsan al-taqāsim fī maʿrefat al-aqālim*, tr. Basil Collins as *The Best Divisions for Knowledge of the Regions*, Reading, UK, 1994.

Enrico Morano, “A Working Catalogue of the Berlin Sogdian Fragments in Manichaean Script,” in Maria Macuch, Mauro Maggi, and Werner Sundermann, eds., *Iranian Languages and Texts from Iran and Turan: Ronald E. Emmerick Memorial Volume*, Wiesbaden, 2007, pp. 239-71.

Takao Moriyasu, “Uigurumoji Shinkou” (A New Study of the Uighur Script), in *Touhougakkai Souritsu Gojjūshūnen Touhougakuronjū/Tohogaku Ronshu. Eastern Studies Fiftieth Anniversary Volume*, Tokyo, 1997, pp. 1226-38.

Takao Moriyasu and Ayudai Ochiru, eds., *Mongorukoku Genson Iseki Hibun Chousa Kenkyū Houkoku/Provisional Report of Researches on Historical Sites and Inscriptions in Mongolia from 1996 to 1998*, Osaka, 1999.

F. W. K. Müller, *Handschriften-Reste in Estrangelo-Schrift aus Turfan, Chinesisch-Turkestan II*, Berlin, 1904.

Idem, *Uigurica I.2*, Abhandlungen der Preußischen Akademie der Wissenschaften, Phil.-hist. Klasse, 1908.

Elio Provasi, “Sogdian Lexicography,” in Carlo G. Cereti and Mauro Maggi, eds., *Middle Iranian Lexicography. Proceedings of the Conference Held in Rome, 9-11 April 2001*, *Orientalia Romana* 8, Rome, 2005, pp. 101-46.

Idem, “Sanskrit and Chinese in Sogdian Garb. The Transcription of Indic Proper Names in the Sogdian Buddhist Texts,” in Mauro Maggi et al., eds., *Buddhism among the Iranian Peoples of Central Asia*, Vienna, 2013, pp. 191-308.

Christiane Reck, *Berliner Turfanfragmente manichäischen Inhalts in soghdischer Schrift*, Stuttgart, 2006.

Idem, *Berliner Turfanfragmente buddhistischen Inhalts in soghdischer Schrift*, Stuttgart, 2016.

Hans Reichelt, *Die soghdischen Handschriftenreste des Britischen Museums, Bd. 2 Die nicht-buddhistischen Texte*, Heidelberg, 1931.

Nicholas Sims-Williams, “The Sogdian Fragments of the British Library,” *Indo-Iranian Journal* 18, 1976, pp. 43-82.



- Idem, review of MacKenzie, 1976, *Indo-Iranian Journal* 20, 1978, pp. 256-60.
- Idem, "On the Plural and Dual in Sogdian," *BSOAS* 42/2, 1979, pp. 337-46.
- Idem, "The Sogdian Sound-system and the Origins of the Uyghur Script," *JA* 269, 1981a, pp. 347-60.
- Idem, "Some Sogdian Denominal Abstract Suffixes," *Acta Orientalia* 42, 1981b, pp. 11-19.
- Idem, "The Double System of Nominal Inflexion in Sogdian," *Transactions of the Philological Society*, 1982, pp. 67-76.
- Idem, "Indian Elements in Parthian and Sogdian," in Klaus Röhrborn and Wolfgang Veenker, eds., *Sprachen des Buddhismus in Zentralasien. Vorträge des Hamburger Symposions vom 2. Juli bis 5. Juli 1981*, Wiesbaden, 1983a, pp. 132-41.
- Idem, "Chotano-Sogdica (I)," *BSOAS* 46/1, 1983b, pp. 40-51.
- Idem, "The Sogdian 'Rhythmic Law'," in Wojciech Skalmowski and Alois van Tongerloo, eds., *Middle Iranian Studies. Proceedings of the International Symposium Organized by the Katolieke Universiteit Leuven from the 17th to the 20th of May 1982*, Leuven, 1984, pp. 203-15.
- Idem, *The Christian Sogdian Manuscript C2*, Berliner Turfantexte XII, Berlin, 1985a.
- Idem, "A Note on Bactrian Phonology," *BSOAS* 48, 1985b, pp. 111-16.
- Idem, "Syro-Sogdica III: Syriac Elements in Sogdian," in Jacques Duchesne-Guillemin et al., eds., *A Green Leaf. Papers in Honour of Professor Jes P. Asmussen* (Acta Iranica 28), Leiden, 1988, pp. 145-156.
- Idem, "Sogdian," in Rüdiger Schmitt, ed., *Compendium Linguarum Iranicarum*, Wiesbaden, 1989a, pp. 173-92.
- Idem, "Eastern Middle Iranian," in R. Schmitt, ed., *Compendium Linguarum Iranicarum*, Wiesbaden, 1989b, pp. 165-72.
- Idem, *Sogdian and Other Iranian Inscriptions of the Upper Indus*, 2 vols., London, 1989-2.



Idem, “Chotano-Sogdica II: Aspects of the Development of Nominal Morphology in Khotanese and Sogdian,” in Gherardo Gnoli and Antonio Panaino, eds., *Proceedings of the First European Conference of Iranian Studies I*, Rome, 1990, pp. 275-96.

Idem, “The Sogdian Inscriptions of Ladakh,” in Karl Jettmar et al., eds., *Antiquities of Northern Pakistan. Report and Studies 2*, Mainz, 1993, pp. 151-63.

Idem, “The Triple System of Deixis in Sogdian,” *Transactions of the Philological Society* 92/1, 1994, pp. 41-53.

Idem, “The Sogdian Merchants in China and India,” in Alfredo Cadonna and Lionello Lanciotti, eds., *Cina e Iran*, Florence, 1996a, pp. 45-67.

Idem, “On the Historic Present and Injunctive in Sogdian and Choresmian,” *Münchener Studien zur Sprachwissenschaft* 56, 1996b, pp. 173-89.

Idem, “The Sogdian Manuscripts in Brāhmī Script as Evidence for Sogdian Phonology,” in Ronald E. Emmerick et al, eds., *Turfan, Khotan und Dunhuang. Vorträge der Tagung “Annemarie von Gabain und die Turfanforschung” veranstaltet von der Berlin-Brandenburgischen Akademie der Wissenschaften in Berlin (9.-12. 12. 1994)*, Berlin, 1996c, pp. 307-15.

Idem, “The Iranian Inscriptions of Shatial,” *Indologica Tauriensia* 23-24, 2000, pp. 523-41.

Idem, “The Sogdian Ancient Letter II,” in M. G. Schmidt and W. Bisang, eds., *Philologica et Linguistica. Historia, Pluralitas, Universitas. Festschrift für Helmut Humbach zum 80. Geburtstag am 4. Dezember 2001*, Trier, 2001a, pp. 267-80.

Idem, “Middle Persian *padisāy* and Old Persian *vašnā*,” in Ali Ashraf Sadeghi, ed., *Yād-nāma-ye Doktor Aḥmad-e Tafazzoli*, Tehran, 2001b, pp. 59-65.

Idem, “The Parthian Abstract Suffix -yft,” in John H. W. Penny, ed., *Studies in Honour of Anna Morpurgo Davies*, Oxford, 2004, pp. 539-47.

Idem, “Towards a New Edition of the Sogdian Ancient Letters: Ancient Letter 1,” in Étienne de la Vaissière and Éric Trombert, eds., *Les Sogdiens en Chine*, Paris, 2005, pp. 181-93.

Idem, *Bactrian Documents from Northern Afghanistan II: Letters and Buddhists*



Texts, London, 2007.

Idem, “The Sogdian Potentialis,” in Maria Macuch, Mauro Maggi, and Werner Sundermann, eds., *Iranian Languages and Texts from Iran and Turan. Ronald E. Emmerick Memorial Volume*, 2007a, Wiesbaden, pp. 377-86.

Idem, “Sogdian-Turkish Bilingualism and Linguistic Interference in 9th-10th Century Tunhuang,” in Zohra Zaršenās and Vidā Naddāf, eds., *Jašn-nāma-ye doctor Badr-al-Zamān-e Qarib / Papers in Honour of Professor B. Gharib*, Tehran, 2008, pp. 40-51.

Idem, “Christian Literature in Middle Iranian Languages,” in Maria Macuch and Ronald E. Emmerick, eds., *A History of Persian Literature XVII. The Literature of Pre-Islamic Iran*, London, 2009, pp. 266-87.

Idem, *Iranian Manuscripts in Syriac Script in the Berlin Turfan Collection*, Stuttgart, 2012.

Idem, “The Vocative Singular feminine in Sogdian,” in V. Naddaf et al., eds., *Ranj o Ganj. Papers in Honour of Professor Z. Zarshenas*, Tehran, 2013, pp. 87-93.

Idem, *Biblical and Other Christian Sogdian Texts from the Turfan Collection*, BTT XXXII, Turnhout, 2014.

Idem, *The Life of Serapion and Other Christian Sogdian Texts from the Manuscripts E25 and E26 BTT XXXV*, Turnhout, 2015.

Idem, *A Dictionary: Christian Sogdian, Syriac and English*, Wiesbaden, 2016a.

Idem, “From Aramaic to Manchu: Prehistory, Life and After-life of the Sogdian Script,” in Rong Xinjiang and Luo Feng, eds., *Sogdians in China: New Evidence in Archaeological Finds and Unearthed Texts*, Beijing, 2016b, pp. 414-421.

Nicholas Sims-Williams, Frantz Grenet, and Étienne de la Vaissière, “The Sogdian Ancient Letter V,” *Bulletin of the Asia Institute* 12, 1998, pp. 91-104.

Nicholas Sims-Williams and Desmond Durkin-Meisterernst, *Dictionary of Manichaean Texts III/2. Dictionary of Manichaean Sogdian and Bactrian*, Turnhout, 2012.

Nicholas Sims-Williams, Frantz Grenet, and Aleksandr Podushkin, “Les



plus anciens monuments de la langue sogdienne: Les inscriptions de Kultobe au Kazakhstan,” *Comptes rendus des séances de l’Académie des Inscriptions et Belles-Lettres*, 2007, pp. 1005-34.

Nicholas Sims-Williams and James Hamilton, *Documents turco-sogdiens du IX^e – X^e siècle de Touen-houang*, London, 1990. English tr. as *Turco-Sogdian Documents from 9th-10th Century Dunuang*, Corpus Inscriptionum Iranicarum Part II, Vol. III/IV, London, 2015.

Nicholas Sims-Williams and Bi Bo, “Sogdian Documents from Khotan, I: Four Economic Documents,” *JAOS* 130/4, 2011, pp. 497-508.

Idea, “Sogdian Documents from Khotan, II: Letters and Miscellaneous Fragments,” *JAOS* 135/2, 2015, pp. 261-282.

Prods Oktor Skjærvø, “Aramaic Scripts for Iranian Languages,” in Peter T. Daniels and William Bright, eds., *The World’s Writing Systems*, New York, 1996, pp. 515-35.

Werner Sundermann, “Manichaean Literature in Iranian Languages,” in M. Macuch and R. E. Emmerick, eds., *A History of Persian Literature XVII. The Literature of Pre-Islamic Iran*, London, 2009, pp. 197-265.

Idem, *Die Rede der Lebendigen Seele. Ein manichäischer Hymnenzyklus in mittelpersischer und soghdischer Sprache*, BTT XXX, Turnhout, 2012.

Idem, “On the Growing Turkicization of the Sogdian Language,” in E. A. Csato, L. Johanson, A. Rona-Tas, and B. Utas, eds., *Turks and Iranians. Interactions in Language and History*, Wiesbaden, 2016, pp. 29-36.

Dieter Weber, *Die Stellung der sog. Inchoativa im Mitteliranischen*, doctoral diss., Göttingen, 1970.

Idem, “Also sprach Mandri (Bemerkungen zur direkten Rede im Sogdischen),” *Indogermansiche Forschungen* 76, 1971, pp. 77-83.

Antje Wendtland, “Zu <h> in Soghdischen,” *Göttinger Beiträge zur Sprachwissenschaft* 1, 1998, pp. 101-10.

Idem, *Die Entwicklung von Demonstrativpronomen zu Artikeln im Sogdischen*, Wiesbaden, 2011.



- Ilya Yakubovitch, “Mugh 1.I Revisited,” *Studia Iranica* 31/2, 2002, pp. 231-53.
- Idem, “The Syntactic Evolution of Aramaic ZY in Sogdian,” *Studia Iranica* 34/2, 2005, pp. 199-230.
- Idem, “Marriage Sogdian Style,” in Heiner Eichner et al., eds., *Iranistik in Europa — Gestern, Heute, Morgen*, Vienna, 2006, pp. 307-44.
- Yutaka Yoshida, “On the Sogdian infinitives,” *Journal of Asian and African Studies* 18, 1979, pp. 181-95.
- Idem, “Kirisutokyou Sogudogono Hougenni tsuite” (On the dialectology of Christian Sogdian), *Oriente* 23/1, 1980, pp. 83-93.
- Idem, “Karabarugasun Hibunno Dogudogobanni tsuite” (On the Sogdian version of the Kara Balgasun inscription), *Bulletin of the Society of Western and Southern Asiatic Studies, Kyoto University* 28, 1988, pp. 24-52.
- Idem, “Sogudomojide Hyoukisureta Kanjion” (Chinese in Sogdian script), *The Toho Gakuho* 66, 1994, pp. 271-380.
- Idem, “Sogudogo Zatsuroku IV” (Sogdian miscellany IV), *Studies on the Inner Asian Languages* X, 1995, pp. 67-84 with plate III.
- Idem, “The Sogdian Version of the New Xi’an Inscription,” in Étienne de la Vaissière and Éric Trombert, eds., *Les Sogdiens en Chine, Paris*, 2005, pp. 57-72.
- Idem, “Sogudogono Keigoni tsuite” (Honorific and polite expressions in Sogdian), in Satoko Shirai and Masahiro Shogaito, eds., *Philological Studies on Old Central Asian Manuscripts*, Kyoto, 2006, pp. 81-94.
- Idem, “Sogdian,” in Gernot Windfuhr, ed., *The Iranian Languages*, London, 2009a, pp. 279-335.
- Idem, “Buddhist Literature in Sogdian,” in Maria Macuch and Ronald E. Emmerick, eds., *A History of Persian Literature XVII. The Literature of Pre-Islamic Iran*, London, 2009b, pp. 288-329.
- Idem, “Minor Moods in Sogdian,” in Kazuhiko Yoshida and Brent Vine, eds., *East and West. Papers in Indo-European Studies*, Bremen, 2009c, pp. 281-93.
- Idem, “Turco-Sogdian Features,” in Werner Sundermann, Almut Hintze, and



François de Blois, eds., *Exegisti monumenta: Festschrift in Honour of Nicholas Sims-Williams*, Iranica 17, Wiesbaden, 2009d, pp. 571-86.

Idem, “Some New Readings in the Sogdian Version of the Karabalgasun Inscription,” in Mehmet Ölmez et al., eds., *From Ötüken to Istanbul. 1290 Years of Turkish (720-2010)*, Istanbul, 2011, pp. 77-86.

Idem, “Sogudojinto Kodaityurukuzoku tonno Kankeini kansuru Mittsuno Oboegaki” (Three philological notes on the Sogdo-Turkish relationship), *Memoirs of the Faculty of Letters, Kyoto University* 50, 2011a, pp. 1-41.

Idem, “When did Sogdians begin to write vertically?” *Tokyo University Linguistic Papers* 31 (Festschrift for Prof. Hiroshi Kumamoto), 2013, pp. 375-94.

Yutaka Yoshida and Takao Moriyasu, “Kikushi Koshoukoku Jidai Sogudobun Onnadorei Baibaimonnzyo” (A Sogdian sale contract of a female slave from the period of the Gaochang kingdom under the rule of Qu clan), *Studies on the Inner Asian languages* 4, 1988, pp. 1-50.