CHAPTER FIVE

SOGDIAN

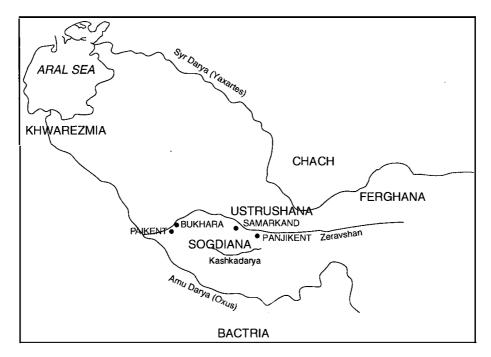*Yutaka Yoshida*

1 INTRODUCTION

I.1 Overview

1.1.1 History

Sogdian is a Middle Iranian language once spoken in Sogdiana. Among the Middle Iranian languages Sogdian is classified as belonging to the North–Eastern group which also includes Khotanese, Tumshuqese, Bactrian, and Choresmian.

The land of Sogdiana is located between the two great rivers of Central Asia, the Amu-Darya and the Syr-Darya (part of modern Uzbekistan and Tajikistan). Sogdiana consisted of several oasis-states which were located along the rivers Zarafshan and Kashka-Darya, the leading cities being Samarqand, Buchara, and Kish (present day Shahr-i Sabz). The eastward movement of the Sogdians led them to colonize Tashkent and an area beyond, and consequently a Chinese pilgrim Xuanzang who went to India via Central Asia in the seventh century referred to the region between Semirechie and Iron Gate, located to the south of Kish, as Suli, i.e. Sogdiana.

The history of Sogdiana is largely obscure. It constituted a satrapy of the Achaemenian Empire, and after the conquest of the empire by Alexander the Great in the fourth century BC, Sogdiana was under the control of neighbouring super powers, such as the Kushans (first to third centuries), the Sasanians (third century), the Kidarites (fourth to fifth centuries?), the Hephtalite (fifth to sixth centuries), the Western Turks (sixth to seventh centuries), and Chinese (seventh to eighth centuries). However, it was able to enjoy a degree of independence until it was conquered by the Arabs in the eighth century. During this period of relative independence the Sogdians played an active role as international traders along the Silk Road between China and the West, with the result that the Sogdian language became a kind of lingua franca in the region between Sogdiana and China, where the Sogdians founded many trade diasporas.

The dominion by the Achaemenian Empire was important for Sogdian in that the chancellery language of the Empire, Aramaic, was introduced to Sogdiana, from which later the Sogdian script developed after the Empire had collapsed and Aramaic ceased to be used by the local scribes. By the end of the sixth century, this Sogdian script came to be written vertically rather than horizontally from right to left, and lines running from left to right.

MAP 5.1 SOGDIA AND ITS NEIGHBOURS

1.1.2 Materials

The Sogdian materials handed down to us reflect the activities played by the Sogdians in history. (On the major publications until 1995, see Gharib 1995: xiii–xxvii.) Except for the very short inscriptions discovered in Kultobe, Kazakhstan (cf. Sims-Williams and Grenet 2006) and the legends found on coins of the early centuries of CE, the earliest substantial materials are the so-called 'Ancient Letters' discovered by Sir Aurel Stein in a watch tower located on the route between Dunhuang and Loulan. They were sent to the addressees staying in Samarqand and oasis cities lying in between by Sogdian merchants and their families resident in China, who reported the news about their activities, and were shown to have been written in the early fourth century (cf. Grenet and Sims-Williams 1987). Numerous rock inscriptions or graffiti discovered in North Pakistan along the ancient route connecting Central Asia and North West India are similar to the Ancient Letters in the ductus and seem to have been inscribed by merchants who came there either directly from Sogdiana or by way of China.

   Their activities of the sixth to seventh centuries are reflected by an epitaph dated 580 CE which was unearthed from a tomb of a Sogdian couple in Xi'an and a sale contract of a female slave dated 639 CE and discovered in Turfan (see Yoshida 2005 and Yoshida et al. 1989). Apparently, during the seventh to eighth centuries when the Tang Dynasty was the most powerful, Sogdian trade activities reached their height, and a great many Sogdians immigrated to China. Some of them were converted to Buddism under the influence of Chinese Buddhists. They produced many Buddhist Sogdian texts based on Chinese prototypes, of which dozens of specimens were discovered in Dunhuang and Turfan.

   The role played by the Sogdians in the propagation of the Manichaeism is not entirely clear. However, a number of Manichaean Sogdian texts were discovered in Turfan, where the Uighurs, who adopted Manichaism in 762/3 CE as their state religion, constituted the

ruling class. Apart from the translations from the Middle Persian and Parthian originals, one finds many works originally composed in Sogdian. Christian monks resident in a monastery of Bulayïq located to the north of the Turfan basin left numerous Christian Sogdian texts based on the Syriac originals. That they were also engaged in trade is inferred from a few Dunhuang texts written by Christians. Apart from the Sogdian script, which was so to speak the Sogdian national script, Manichaean and Syriac scripts were employed respectively to write Manichaean and Christian texts.

Similarly, Brahmi script was applied to write texts translated from Sanskrit (or Tocharian as well?), either of Buddhist or medical contents. However, only a small number of Brahmi fragments have so far been encountered among the treasures unearthed from Central Asia (see Sims-Williams 1996c).

While the bulk of Sogdian materials are from East Turkestan, the land of Sogdiana also yields a handful of manuscripts and short inscriptions. The most famous are some 80 secular documents discovered at Mount Mugh, which represent part of archives of Dhēwaštīč (?–722), the last independent ruler of Penjikent.

While the Sogdians were governed by Turkish peoples, such as Turks and Uighurs, they exercised strong cultural influence upon their rulers; the Sogdian language served as a literary language for them until they began to write their own languages, first in Runic, and later in Uighur script, of which the latter originated from the cursive variant of the Sogdian script. Of the two Sogdian inscriptions unearthed in Mongolia, one, the Bugut Inscription, was erected by Turks around 600 CE and the other, the Karabalgasun Inscription, is trilingual in Sogdian, Chinese and Uighur in Runic script, and commemorates the eighth Khaghan of the Uighur empire (r. 808–822). Another stele which is similar in contents to the Bugut Inscription was discovered on a stone statue unearthed in Mogolküre in Xinjiang.

Very late inscriptions going back to the tenth to eleventh centuries, which have been discovered in Kirghizia, contain several Turkish elements; they may have been written by Turkophones who employed Sogdian as a written language (see Livšic 1996). A few late documents also containing Turkish words and phrases were discovered in Dunhuang and Turfan.

The linguistic variations found among all these materials are not trivial mainly for chronological reasons. Moreover, the bulk of Christian Sogdian texts and a great part of the Manichaean materials are written respectively in Syriac and Manichaean scripts, whose orthographies are free from historical spellings found in texts written in Sogdian script. However, despite the wide area of diffusion of the Sogdian language, very few purely dialectal differences have so far been observed. On this point see the section 7.2, Dialects and variations.

1.2 Scripts, orthography, and phonetic values

The three scripts (Sogdian, Manichaean, and Syriac) derive ultimately from the Aramaic and thus have many features in common, the most salient being that they are so-called consonant scripts, and that vowel phonemes are not always sufficiently distinguished. Inevitably some uncertainties surround the problem of the Sogdian vowel phonemes, the situation not being radically improved by the recent discovery of a few texts written in Brahmi script (see Sims-Williams 1996c). In this connection an important advantage of Syriac script is to be mentioned; vocalic points are sometimes employed to distinguish vowels of different qualities: y: = [ĕ], y: = [ĭ], ẇ = [ŏ], ẇ = [ŭ], and Ç = [Ca]. In the texts written in Sogdian script, which evolved from the Achemenid chancery Aramaic, one

TABLE 5.1: SOGDIAN SCRIPTS

Aramaic	"Ancient Letters"	"Sogdian" (Buddhist) Sogdian Script[1]			"Manichean" Sogdian Script[2]			"Syriac" (Christian) Sogdian Script		
			< >	[ ]		< >	[ ]		< >	[ ]
'			'	ā, a (ə, i)		'	ā, a (ɔ, i)		'	ā, a (ə, i)
b			β	β, f		b	b		b	β, b
			(β =) f	f		(b̈ =) β	β			
g			γ	γ		g	g		g	g
						γ	γ			
d			D[3]	-		d	d		d	δ
h			h[4]	-ă, Ø		ẖ[4]	Ø		h	h
w			w	w, ŏ, ŭ		w	w, ŏ, ŭ		w	w, ŏ, ŭ
z			z	z, ž		z	z		z	z
			(z, ẓ =) ž[5]	ž		j	ž, ǰ		ž	ž
ḥ			x	x, h		h	h		ẖ	h(?)
ṭ			ṭ[7]	-		ṭ	t (d)		t	t (d)
y			y	y, č, ī		y	y, č, ī		y	y, č, ī

k			k	k (g)		k	k		k	k⁶
						(k̇ =) x	x		(k̇ =) x	x̱
l			δ/L³	δ, θ, l		l	l		l	l
						δ	δ, θ			
m			m	m		m	m		m	m
n			n	n, ṃ		n	n, ṃ		n	n, ṃ
s			s	s		s	s		s	s
ʿ			ʿ³			ʿ	ɨ		ʿ	γ
p			p	p (b), f		p	p (b)		p	p (b)
						f	(ṗ =) f		f	f
ṣ			c	č (ǰ)		c	č (ǰ)		c	č (ǰ), ts
q			q⁷	–		q	k		q	k (g)
r			r	r, ʾ, l		r	r, ʿ		r	r, ʾ
š			š	š		š	š		š	š
t			t	t (d)		t	t (d)		θ	0

finds a certain number of ideograms, i.e. Aramaic word forms pronounced with their Sogdian equivalents (see also Chapter 4 on Middle West Iranian). For example, a word for 'ten thousand' is often spelled as RYPW (cf. Aramaic rbw = ribbō) and was pro- nounced as βrēwar. That it was pronounced as a Sogdian word is proved by a phonetic complement which appears in its cardinal Aramaic form with the Sogdian ordinal suffix (-myk): (RYPWmyk =) *βrywrmyk, i.e. βrēwar-mīk '10,000th'. (In modern trans- literations one distinguishes ideograms by Roman capital letters.) Below is the chart of Sogdian, Manichaean, and Syriac scripts with their transliteration and phonetic values, which will be discussed in the following section.

1   Of the two major varieties, formal and cursive scripts, the former as attested in the manuscript of the *Vessantara Jātaka* is reproduced.
2   The frequent doubling of the letter y, w, and δ, and ṭ has no phonetic significance.
3   Only used in ideograms.
4   Only used in word-final position.
5   Usage varies: some scribes employ the diacritic point(s) to distinguish z (= [z] or [ž]) from n, others to distinguish [ž] from [z].
6   Only in knθ, kθ 'city' and in Syriac words.
7   Not used.

2 PHONOLOGY

In the following, M and C indicate sources in Manichaean script (without distiction of t and ṭ) and Christian Syriac script, respectively. Otherwise, spellings cited are those in Sogdian script (if necessary indicated by S). Light stems, on which see below, are differentiated from heavy with a hyphen (-).

2.1 Inventory and distribution

2.1.1 Vowels and diphthongs

2.1.1.1 Vowel system

Given that the three scripts, Sogdian, Manichaean, and Syriac, ultimately derive from the Aramaic consonant script, there remain uncertainties in the determination of the Sogdian vowel phonemes, which are inconsistently indicated.

TABLE 5.2: VOWELS AND DIPHTHONGS

	front	central	back	Diphthongs
high	ī, i	(i)	u, ū	
mid	ē, e	(ə)	o, ō	
low	a	ā		āi āu; V + m̨; V + əʳ

There is phonemic length distinction, which is not overtly marked except for the low pair, and is otherwise inferred etymologically as well as by means of the 'rhythmic law' (see below): βay- (βγ-) 'god', βāy (β'γ) 'garden, farm'; but *fneš-* (C fnyš-) 'be deceived',

fnēš (C fnyš) 'deceive'; *witar-* (wytr-) 'go', *wītar* (wytr) 'went (3s imperf.)'; *but-* (pwt-) 'Buddha', *pūtē* (pwt'k) 'rotten'; *roxšn-* (rwxšn-) 'light, bright', *rōyn* (rwγn) 'oil'.

Naturally, long and short vowels are likely to be different in terms of quality as well. Thus *ā* is a back vowel in contrast with *a* which is front. This is clear in view of the Uighur orthography which is based on the Sogdian spelling conventions; thus in Uighur a back vowel *a* is expressed by initial (''-) and medial (-'-), both being the notations for *ā* in Sogdian, whereas the front counterpart *ä* is represented by those standing for *a* in Sogdian (see Sims-Williams 1981: 358).

i and *ə* are allophones of a phoneme *a* mainly appearing in the initial position where *a* never occurs: *əpsāk* ('ps'k) 'garland'. Compare also the fact that foreign words beginning with *a-* are usually transcribed with (''-), i.e. *ā-* rather than ('-), i.e. *ə-*, e.g. (''swr) for Skt. *asura*.

The high central vowel *i* is restricted to the position before *sp-*, *st-*, and *sn-* and is transcribed by the letter ' ('ain) in Manichaean script: *isptyāk* (M 'spty'k) 'completion'. When they are unaccented, short vowels *i*, *a*, and *u* seem to have been reduced to *ə*. The alternation of (-y-), (-w-), and (zero) encountered in the spellings seems to reflect this situation: e.g. *žət-* (M jt- ~ jyt-) 'strike' (past stem < OIr. *jata-*); *δəštē* (M δyštyy ~ S δšt'y) 'built' (< OIr. *dištaka-*); *ōsɔγtē* ('ws(')γty ~ 'ws(')wγtk) 'pure' (< OIr. *awa-suxtaka-*).

2.1.1.2 Rhotacised vowels

Apart from the above-mentioned simple vowels, Sogdian possesses three rhotacised vowels, *ə'*, *i'*, *u'* which are counted as short vowels: *mə'γ-* (mrγ-) 'bird', *ki'm-* (kyrm-) 'snake', and *pu'n-* (pwrn-) 'full'. This element constitutes the second element of diphthongs: *mary = maə'γ* (mrγ) 'forest', *zern = zeə'n* (zyrn) 'gold'.

2.1.1.3 Diphthongs

As the second member of diphthongs Sogdian also possesses a nasal element *m̥*, which probably had moraic status: *kam̥θ* (knδh) 'city', *pam̥j* (pnc) 'five'.

The long diphthongs *āi* and *āu* are also known: *āikūn* (''ykwn) 'forever', *āurāθē* (''wr'δ'k) 'fellow-traveller', while Old Iranian **ai* and **au* have become *ē* and *ō* respectively.

2.1.2 Consonants

2.1.2.1 Consonantal system

Sogdian has the following consonants; those in parentheses are allophones or marginal phonemes.

TABLE 5.3: CONSONANTS

	labial	dental	alveolar	palatal	velar/glottal
Plosive/affr.	*p (b)*		*t, (d) (ts)*	*č (ǰ)*	*k (g)*
Fricative	*f β*	*θ δ*	*s z*	*š ž*	*x γ*
Nasal	*m*		*n*		*(ŋ)*
Liquid/Glide	*w*		*r (l)*	*y*	*(h)*

The typologically marked system with voice opposition found only with fricatives is due to the sound change in which the voiced plosives and affricate *b, *d, *ǰ, *g have become fricatives β, δ, ž, and γ even in intial position: βar- (βr-) 'bring', δār (δ'r) 'to hold', žār (C ž'r) 'poison', γar- (γr-) 'mountain'. On the other hand, due to the Sogdian conservatism, the voiceless plosives and affricate *p, *t, *č, *k even after a vowel are preserved: āp (' 'p) 'water', wāt (w't) 'wind', wyāk (wy'k) 'place', wāč (w'c) 'send', 'release'.

The voiced counterpart b, d, ǰ, g are found only after nasalized vowels, i.e. m̥ zamb (M zmb) 'coast, river bank', βaṃdē (M βndy) 'slave', paṃǰmīk (M pnjmyk) 'fifth', saṃg (M sng) 'stone'.

2.1.2.2 Marginal l, h, ts, and ŋ;

l and h are marginal phonemes only found in loan words and foreign forms: šlōk (šl'wk ~ šr'wk) < Skt. śloka 'verse'; rahaṃd (M rhnd) < Skt. arhant 'arhant'. Only sporadically do l and h appear in original Sogdian words as a result of dissimilation: wilarz (C wlrz ~ S wyr'rz) 'tremble' (< OIr. *wi-rarz-, cf. unreduplicated form attested in Khotanese rrīyz- 'tremble'), hunax ~ xunax reflected in (M hwnx ~ xwnx) 'that'.

The marginal status of ts may be inferred from the fact that the Christian Sogdian form (mc') 'hither' corresponds to (mrts'r) in Sogdian script; thus Syriac tzaddi (c) can represents č as well as ts, so that ts is treated as a single segment. Compare also (M pncts) and (C pncc) both representing panč(a)ts 'fifteen'.

ŋ is sometimes represented by the spelling (-nn-) in Manichaean script: sarθaŋg (srδnng beside srδng), i.e. saə'θṃg 'leader'.

2.1.3 Syllable structure and clusters

Due to the Sogdian orthography with consonant scripts it is not always easy to know the syllabic structures. Etymological consideration and spellings lead us to assume that Sogdian allows consonant clusters consisting of a considerable number of consonants: žuxšk- (C žwxšq-) 'disciple', sfrīn (M sfryn) 'create', xēpθtī (C xypθ-t-y) 'one's own (pl. obl.)'.

The pronunciation of clusters is occasionally eased by prothetic vowel or by metathesis: psāk (M ps'k) ~ əpsāk (M 'ps'k) 'garland', žuxšk- (C žwxšq-) ~ žxušk- (žxwšk-, δrxwšk-) 'disciple', etc.

Combinations of voiced fricative and voiceless plosive or affricate are characteristic of Sogdian, paδk- (pδk-) 'law', əβta ('βt') 'seven'. In view of the similar clusters in Ossetic, one may assume that the second member was pronounced as a weak devoiced plosive, i.e. əβd̥a.

2.2 Non-segmental features and rhythmic law

(In the following stress is indicated by an underlined vowel.) The position of the stress accent is governed by the so-called 'rhythmic law' (see Sims-Williams 1984) as follows:

1. A syllable containing either a long vowel or diphthong is counted as heavy while a syllable consisting only of a short vowel is light.
2. Stress falls on a stem if it contains a heavy syllable, i.e. heavy stem, whereas it falls on the suffix or ending in the case of light stems that consist only of light syllable(s). Unstressed endings and suffixes suffer reduction or loss, so that the patterns of

conjugation and declension of heavy stems are totally different from those of light stems; thus:

heavy stem: *mēθ* (C myθ) 'day (dir. sg.)', *mēθ-ī* (myθy) (obl. sg.);

light stem: *ram-i* (rmy) 'people (nom. sg.), *ram-u* (rmw) (acc. sg.), *ram-ya* (rmy') (loc. sg.), etc.

Enclitic and proclitic forms, even if they consist of light syllables, do not bear stress and lose their endings: *βay-a* (βγ') vs. -*βay* (-βγ) 'o lord!' (encl. voc.); *əsti* ('sty) vs. *əst* ('st) 'is' (encl.);

Light stems ending with -*r*, -*w*, and -*y* sometimes behave as heavy when they are followed by endings beginning with a consonant: *βar-t* (βrt) ~ *βar-ti* (βrty) 'bears' (3 sg. pres.).

2.3 Alternations and combinatory phenomena

A considerable amount of spelling variations are observed throughout the texts; some are simply due to the difference between historical spellings and those more or less reflecting the actual pronunciation: *(ə)xšēθ* ('xš'y"δ vs. M xšyδ) 'king' (<Olr. *xšāyaθiya*).

Most conspicuous are the differences observed between Christian Sogdian forms in Syriac script and those in Sogdian script. Prothetic vowels are virtually non-existent in the former while they abound in the latter: *zβāk* (C zb'q) 'tongue' vs. *əzβāk* ('zβ'k). Another peculiar difference between the two groups is the existence of *m̥* in the latter and its loss in the former before continuants: *kam̥θ* (knδh) 'town' vs. *kaθ* (C qθ), *am̥yām* (S, M 'ny'm) 'end' vs. *ayām* (C 'y'm), but compare *βam̥tē* (S βnt'k, M βndy, C bnty) 'slave'. In late 'demotic' texts one sometimes finds the tendency to voice *t* and *p* after a voiced sound: *āb* (C 'b, cf. M "p) 'water', *mardəxmē* (C mrdxmy, cf. M mrtxmy) 'man', *saydya* (C sydy', cf. M syty') 'on the . . . th day'.

Similarly, post-vocalic *r* = *ə*' is often lost in the Christian texts: S *zeə'n (zyrn)* 'gold' vs. C *zen* (zyn); S, M *sāə'* (s'r) vs. C *sā* (s').

The loss of *m̥* and *ə'* does not change the rhythmic status of the stems. Accordingly, Christian Sogdian *kaθ* and *zen* behave as heavy stems.

Metathesis of *u* or *w*, both progressive and regressive, is so common that one almost always finds alternative forms: *δwyt-* (δwγt-) ~ *δyut-* (δγwt-) 'daughter', *suγδīk* (swγδyk) ~ *syuδīk* (sywδyk) 'Sogdian', *yāθuk* (γ'δwk) ~ *ywāθk* (γw'δk) 'throne', compare Olr. **gāθu-*. This metathesis affects even loanwords: *samutr-* ~ *sumtar-* (smwtr- ~ swmtr- < Skt. *samudra*) 'ocean'.

The cluster *čt* often becomes *št*: *sāčt* (s'ct) ~ *sāšt* (s'št) 'it is necessary', *βaypūrēšt* (< *βaypūrēč* with the plural ending -*t*) 'divine virgins'.

3 MORPHOLOGY

From here on, the transcription is somewhat simplified; thus *n* or *m* for *m̥*, *r* for *ə'*. Also, voiced plosive allophones after *m̥* are transcribed as *p*, *t*, *k* or *b*, *d*, *g* in conformity with spellings: e.g. *βantē* (S βnt'k) or *βandē* (M βndy) for phonemic *βam̥tē*. Occasionally transliterated forms are given, somewhat inconsistently, for the sake of clarity. Moreover, the inflectional tables are idealized in that attested endings are added to sample stems regardless of whether the forms in question are actually attested or not.

3.1 Nominal morphology

The inflections of Sogdian nouns and adjectives are identical, and the following

description also applies to adjectives, except that the latter lacks the numerative, on which see below. On the Sogdian inflection see Sims-Williams 1982 and 1990.

The Old Iranian distinction of three genders (masculine, faminine, and neuter) has been preserved, although the survival of the neuter is marginal as many old neuter nouns have shifted to masculine or feminine. Similarly, the distinction of three numbers has been preserved; however, the old dual forms have come to be used in the position immediately following a numeral, where they occur not only after the number two but also higher numbers, and thus developed in the special form called 'numerative' (NUM), and discovered by Sims-Williams (1979).

The case distinctions, largely reflecting Old Iranian, are: nominative (NOM, N), accusative (ACC, A), genitive-dative (G–D), locative (LOC), instrumental-ablative (I–A), and vocative (VOC), with light stems, which are reduced to the opposition of direct (DIR) vs. oblique (OBL) with heavy stems.

3.1.1 Stem classes and declensions

Sogdian nouns are classified into several stems. Apart from the distinction between light and heavy (hereafter abbreviated as LS and HS repectively), a few light stems ending with -*u* (-w) inflect differently from ordinal light stems.

A considerable number of stems go back to forms extended by the suffixes masc. *-aka* and fem. *-ākā*, and are referred to conventionally as the *aka*-stem and the *ākā*-stem. In Sogdian script, these two stems are often written with historical spellings, (-(')k) and (-'kh), respectively. The feminine counterpart of the masculine stems ending in -*ē* (< *-aka*) usually corresponds with the adjectival feminine stems ending in -*č*: masc. *spt-ē* (C spty) 'complete', fem. *spč-* < *spt-č-* (C spc-).

There are also a few indeclinable nouns ending in -*ī* (-y): *martī* (mrty) 'man'.

3.1.1.1 Light stems

The following are paradigms of light stem nouns masc. *ram-* 'people', neut. *βaγn-* 'temple, altar', and fem. *wan-* 'tree'.

The regular plural forms are marked by the ending -*t*-. These plural stems are treated as feminine singular; that is, the plural form of light stems like *ram-* have the same inflection as feminine light stems like *wan-*.

TABLE 5.4: DECLENSION OF LIGHT STEMS

	Singular			Plural	
	'people' masc.	'temple' neut.	'tree' fem.	'people' masc.	'tree' fem.
	ram-	*βaγn-*	*wan-*	*ram-t-* (pl.)	*wan-t-* (pl.)
NOM	*ram-i*	*βaγn-u*	*wan-a*	*ram-t-a*	*wan-t-a*
ACC.	*ram-u*	*βaγn-u*	*wan-a*	*ram-t-a*	*wan-t-a*
G–D	*ram-e*	*βaγn-e*	*wan-ya*	*ram-t-ya*	*wan-t-ya*
LOC	*ram-ya*	*βaγn-ya*	*wan-ya*	*ram-t-ya*	*wan-t-ya*
I–A	*ram-a*	*βaγn-a*	*wan-ya*	*ram-t-ya*	*wan-t-ya*
VOC	*ram-a*		*wan-e*	*ram-t-e*	*wan-t-e*
NUM, N–A	*ram-a*	*βaγn-e*	*wan-e*		

For a masculine *u*-stem noun, note nom.-acc. *mayu* (mγw) 'magus', gen.-dat. *mayw-e* (mγwy).

3.1.1.2 Heavy stems

As indicated above, the declension of heavy stems does not differentiate masculine from feminine and distinguishes only direct and oblique cases. It is to be noted that in actual texts, especially those in Sogdian script, heavy stems often take the endings of light stems by analogy and possibly because of the clarity and salience of light stem inflection for the syntactic function.

TABLE 5.5: DECLENSION OF HEAVY STEMS

	masc. *mēθ* 'day'	pl.	fem. *žwān* 'life'	pl.
DIR	*mēθ*	*mēθ-t*	*žwān*	*žwān-t*
OBL	*mēθ-ī*	*mēθ-t-ī*	*žwān-ī*	*žwān-t-ī*
VOC	*mēθ(a)*	*mēθ-t-e*		*žwān-t-e*

3.1.1.3 Irregular plural forms, *-ar-t*, *-īš-t*

The plural forms of *βrāt* 'brother' and *δuyt-* 'daughter' have the formant *-art-*: *βrāt-ar-t*, *δuyt-ar-t*. Also, some animate nouns, especially light stems, have the plural markers dir. *-īš-t*, obl. *-īš-t-ī*: *βay-* 'god', pl. *βay-īš-t-*.

3.1.1.4 *aka-* and *ākā-*stems

Historically, intervocal *-k-* was lost, and the resulting hiatus was later contracted (see Sims-Williams 1990: 286–291); thus, nom. masc. sg. **-aki* > *-*a'i* > *-ē*; similarly: acc. **-aku* > **-a'u* > *-ō*. In further development, while in one Christian Sogdian manuscript C2 the original case distinctions are well preserved after the contraction, in all the other texts the ending *-ē* found in the nom., gen.-dat., and loc. sg. was generalized, replacing acc. *-ō*, and instr.-abl., voc. *-ā*. The plural ending *-t* is suffixed to the nominative form, thus *-ē-t*, and inflects as the heavy stem.

The contraction in the feminine *ākā-* stems similarly resulted in minimal distinctions. The following table shows the declension patterns of masc. *martǝxmē* 'man' in C2 and in other texts, and of fem. *xānā* 'house'.

TABLE 5.6: DECLENSION OF *aka-* AND *ākā-* STEMS

	masc. *martǝxm-ē* 'man'				fem. *xānā* 'house'		
	C2	other texts	pl.		sg.	pl.	
NOM	*martǝxmē*	*martǝxmē*	DIR	*martǝxmē-t*	DIR	*xānā*	*xānē-t*
ACC	*martǝxmō*	*martǝxmē*	OBL	*martǝxmē-t-ī*	OBL	*xānē*	*xānē-t-ī*
G-D	*martǝxmē*	*martǝxmē*	VOC	*martǝxmē-t-e*	VOC		*xānē-t-e*
LOC	*martǝxmē*	*martǝxmē*					
I-A	*martǝxmā*	*martǝxmē*					
VOC	*martǝxmā*	*martǝxmē*					
NUM, N-A	*martǝxmā*	*martǝxmē/ā*					

3.1.1.5 Indeclinables and archaic plural -ān

Indeclinable nouns do not inflect in the singular, but have the binary case distinctions in the plural ending: sg. dir., obl. *martī* 'man', pl. dir. *martī-t*, obl. *martī-tī*. Some nouns take the old genitive plural ending -ān (< OIr. *-ānām): *βay -ān* 'gods' (LS), *pīδ-ān* 'elephants' (HS), *martəxm-ān* (< *-a'ān*, *aka*-stem). This ending is restricted to stereotyped phrases such as *βay-ān βaxtam* 'godliest of gods' = Skt. *devātideva*.

3.1.2 Adjectives

3.1.2.1 Degree

The productive suffix of the comparative is -*(i)star*: *murzək-istar* 'shorter', *yawāncīk-star* 'more necessary'. Somewhat obsolete is the suffix *-tar*, which is not suffixed to derived stems or compounded forms. Thus we have *δūr-tar* 'farther', *namr-tar-* 'sweeter' but *mand-yrβāk-star* 'more stupid', *pačxūδ-yōnē-star* 'more despicable', etc. Irregular forms are: *mazēx* 'big, great' > *masyātar*, *yarf* 'many, much' > *fyātar*, *šir-* 'good' > *šyātar*, *kaβn-* 'little, few' > *kampī*, etc.

Old superlative forms with the suffix *-tam* are attested. They are either restricted to stereotype expressions, as in *βay-ān βax-tam* 'devātideva', or lose their superlative meaning and are treated as simple adjectives (to which secondary suffixes are added: *əskā* 'high' ~ *əskā-tam-čīk* 'highest'.

An alternative way is preposing *ēw* 'one' or *āδparm* 'whatsoever' to the comparative form, e.g. *ēw əxšnak-istar* 'the most excellent', *āδparm fratar* 'the best'.

3.1.2.2 Elative

Sogdian possesses special elative formation with the meaning 'so much ~, very ~'. It is formed by means of the exclamatory particles *čā-* 'how' and *wā-* 'so' ~ *wat-* (before *s-*), together with the suffixes *-t*, *-(ə)st*, in various combinations: *čā-əβižāxuk-st* 'so painful', *wā-zārī* 'so miserable', *wā-frāk-t* 'so early in the morning', *wat-spēt-t* 'so white'.

3.1.3 Pronouns and deixis

3.1.3.1 Personal pronouns

The 1s and 2s personal pronouns distinguish direct and oblique. Independent forms of the 3rd person are provided by weak demonstratives (see below):

TABLE 5.7: PERSONAL PRONOUNS

	1s	2s	1p	2p	3s	3p
DIR	*(ə)zu*	*tayu*	*māx*	*(ə)šmāx*	= dem. pron.	
OBL	*mana*	*tawa*	*māx*	*(ə)šmāx*	= dem. pron.	
ENCL	*-mīl-m*	*-fī-f*	*-man*	*-fan*	*-šu, -šī, -š*	*-šan*

Some texts show case distinction in enclitic forms; thus 2s *-f* for acc. and inst.-abl. and *-t(ī)* for gen. -dat. 3s *-šu* is chiefly used for acc. and *-šī* for all cases including acc.

When they are dependent on adpositional elements which are also enclitic (-č 'from', -δ 'with', -t 'to'), forms -m, -f, and -š appear: rti-šč (rty-šc) 'and from him', etc. The 1s and 2s have also fused prepositional case forms: 1s tāmā and 2s tāfā (S t'β'kh) functioning as acc. (t- obsolete ət(ā) 'to'), and similarly čāmā, čāfā (č- 'from'), δāmā, δāfā (δ- 'with'), parāmā, parāfā (par- 'on, by'). (See also Adpositions, section 3.1.4).

3.1.3.2 Reflexive and reciprocal forms

A feminine noun γrīw serves as the reflexive pronoun referring to the subject, while it also retains its original meaning 'body':

wāyōnč īnč   kū əyrīw   δārāt
such   woman to self   hold.SUBJ.3s
'one should keep such a woman with oneself'.

Reciprocal sense is conveyed by ēw 'one' in combination with the inflected form of δəβty- 'second' or by əny- 'other' combined with another əny-. xaδanyu 'each other' enphasizes the meaning: ēw δəβtya xaδanyu nē γərβant 'they do not know each other'.

3.1.3.3 Articles, demonstrative pronouns and adverbials

Sogdian distinguishes two kinds of demonstratives: (1) one employed as the article and 3rd person pronouns (weak demonstratives), and (2) the other extended forms functioning as proper demonstratives (strong demonstratives). The latter forms are extended from the former by means of several elements.

Sogdian distinguishes three foci of deixis, each consisting of a direct and oblique base: (1) y-/m- 'this (with me)', (2) š-/t- 'that (with you)', and (3) x-/w- 'that (with him)' (see Sims-Williams 1994). The three foci are here referred to by Ich-, Du-, and Er- deixis. Strong demonstratives are extended either with -n- or -ēδ. There are adverbial forms based on the demonstrative bases. The following tabulates the declensions of some attested samples with their typical uses.

TABLE 5.8: ARTICLES AND PERSONAL PRONOUNS

Ich-Deixis

	masc.	fem.	pl.
NOM	yu (yw)	yu (yw)	= f.
ACC	(ə)mu ((')mw)	əma('mh)	= f.
		əmu ('mw)	
G-D	əmen ('myn)	əmī ('my)	mēšan (myšn)
LOC	əmya ('my')	əmya	
	əmī ('my)	əmī	

Du-Deixis

	masc.	fem.	pl.
NOM	(ə)šu (((')šw)	ša (š')	= f.
ACC	ətu ('tw)	əta ('th)	= f.

(On (š'), see Yoshida 2000: 82.)

Er-Deixis

	masc.	fem.	pl.
NOM	*(ə)xu* (('')xw)	*xa* (xh, x')	= f.
ACC	*ō* ('w)	*wa* (w')	= f.
		ō ('w)	= f.
G-D	*wənī* (w(y)ny)	*wya* (wy')	*wēšan* (wyšn)
	wənī ('wyn)	*əwī* ('wy)	
LOC	*wya* (wy')	*wya*	
	əwī ('wy)	*əwī*	

TABLE 5.9: DEMONSTRATIVES

n-Extension
Ich-Deixis

	masc.	fem.	pl.
NOM	*yunē* (ywn'k)	*yānā* (y'n'kh)	*yānt* (C y'nt)
	ēnē ('yn'k)		
ACC	*munō* (mwn'kw)	*mānā* (m'n'kh)	*mānt* (C m'nt)
G-D	*nimant(ī)* (nym'nt(y))		
LOC	*əmyamant(ī)* ('my'mnt(y))		

Du-Deixis

	masc.	fem.	pl.
NOM	*šunē* (šwn'k)	—	*šānt* (š'ntt)
ACC	*tunō* (twnkw)	—	—

(On *šunē*, cf. Yoshida 2000: 81.)

Er-Deixis

	masc.	fem.	pl.
NOM	*xunē/xunax* (xwn'k/xwn'x)	*xānā* (x'n'kh)	*xānt* (C x'nt)
ACC	*ōnō/wānō* ('wn'kw/w'nw)	*wānā* (w'n'kh)	*wānt* (C w'nt)
G-D	*(wə)niwant(ī)* ((w)nyw'nt(y))		
LOC	*wyawant* (wy'wnt)		

ēδ-Extension

	Ich-Deixis	Du-Deixis	Er-Deixis
DIR	*ēδ* (' yδ)	—	*xēδ* (xyδ)
OBL	*mēδ* (myδ)	*tēδ* (tyδ)	*wāδ* (wyδ)

TABLE 5.10: DEMONSTRATIVE ADVERBS

	Ich-Deixis	Du-Deixis	Er-Deixis	Interrogative
"here", etc.	*maδe* (mδy)	*taδe* (tδy)	*waδe* (wδy)	
"here", etc.	*maδēδ* (mδyδ)	*taδēδ* (tδyδ)	*waδēδ* (wδyδ)	
"here", etc.	*mar0* (mrδ)	*tar0* (trδ)	*ōr0* ('wrδ)	*kurθ* (kwrδ)
"hither", etc.	*martsār* (mrts'r)	*tartsār* (trts'r)	*ōrtsār* ('wrts'r)	*kurtsār* (kwrts'r)

In Manichaean and Christian texts, the gen.-dat. plurals *mēšan* and *wēšan* provided with the secondary plural ending *-t* are also employed: *mēšan-d/mēšan-d-ī*, *wēšan-d/ wēšan-d-ī*.

These articles, just before they were lost in the latest stage of Sogdian, came to be proclitic and prefixed to nouns: *yi-mān-ī* (C y-m'ny; *yi* < *wya*, *əwī*, loc. masc. of Er-deixis) 'in the mind'.

The articles (weak demonstratives) themselves are occasionally preposed to the extended deictics, e.g. *xu xunax γrīw* 'that body'.

The adjective *wisp-* 'all' occasionally takes pronominal endings: G-D *wisp-ne*, I-A *wisp-na* (cf. *čan wispnā-č* 'from all' and *δan wispnā-δ* 'with all'), PL. NOM *wisp-e*, PL. G-D *wispēšan*.

The functional distinction between simple forms as articles, and the extended forms as demonstratives can be shown by their use in translations. Thus, of the 70 instances of simple forms in lines 1–88 of Pelliot Sogdien 5, the short Buddhist text translated from the Chinese *Dīrghanakha-sūtra*, only one case corresponds to the Chinese *qi* 'that'. In turn, of the 32 instances of extended forms, all but three render a demonstrative.

3.1.3.4 Interrogative, relative, and indefinite pronouns

In Sogdian every interrogative pronoun or adverb can be used as relative pronoun or adverb. The following interrogative pronouns are known in Sogdian: dir. *(ə)ke* 'who?' ((')ky), obl. *(ə)kya* ((')ky'), inst.-abl. *kanāč*, *čakanāč*, or *čakanā* (kn'c, ckn'c, ckn'); *(ə)ču* 'what?' ((')cw); *katār* 'which' (kt'r); *katām* 'which' (kt'm).

While *(ə)ču* is exclusively for inanimate antecedents, *(ə)ke* is the most commonly used relative pronoun, used for both animate and inanimate antecedents. Since *(ə)ke* is sometimes used also for oblique cases, it may rather be referred to as a relative particle that simply connects relative clauses to main sentences. Some Christian texts employ *kat* (qt) as a relative particle.

The following interrogative adverbs are also used as the relatives. They are: *(ə)kū* 'where' ((')kw), *kūr0* 'where' (kwrδ), *kūtsā(r)* 'to where, whence' (kwrts'r, C qwc'), *kaδa* 'when' (kδ'), *čāf* 'how much' (c'β), *čāfar* 'how much' (c'βr), *čānō* 'how' (c'n'kw, C c'nw).

The following forms serve as indefinite pronouns in Sogdian (see Sims-Williams 1986a): animate *āδē* ("δy), *ēδē* ('yδy) 'someone', and inanimate *ā(δ)č* ("δc, "c), *ē(δ)č* ('yδc, 'yc) 'something'. The corresponding negative forms are *nēδē* (nyδy) and *nēδč* (nyδc) with some variant forms. The negated forms always appear with the negative particle (see section 3.2.4), that is to say, doubly negated: *rti-šī nēδē nē pērt* 'and-him nobody not believes = Nobody believes him'.

3.1.4 Adpositions

In Sogdian both prepositions and postpositions are common.

3.1.4.1 Fused prepositions

The inherited prepositional elements are *par-* 'on, in', *č-* 'from', *δ-* 'with', while the obsolete preposition *(ə)t* ('t) < *(ə)tā* (reflected in the 1s, 2s pronouns *tāmā*, *tāfā* (see section 3.1.3.1) has been functionally replaced by *(ə)kū* 'to, towards', originally the older relative adverb *kū* 'where' (see Sims-Williams 1986b). *wasn* (wsn) 'for the sake of' and *witūr* (wytwr) 'until, up to' are not common but are not obsolate either.

What is peculiar to Classical Sogdian is that old prepositions never appear independently but are always fused with oblique pronominal or demonstrative elements. On this point see also the two ideograms found in the Ancient letters, 'LZK = *paru* and 'NwZK = *kū* 'where' which contain the pronominal element ZK = *-u*.

TABLE 5.11: FUSED PREPOSITIONS

	No Extension		*n*-Extension		*ēδ*-Extension	
	Er-Deixis < -*w*-	Ich-Deixis < -*m*-	Er-Deixis < -*want*	Ich-Deixis < -*mant*	Er-Deixis < -*wēδ*	Ich-Deixis < -*mēδ*
par-	*paru* (prw)	*parəm* (prm)	*pariwant*	*parimant*	*pariwēδ*	*parimēδ*
c-	*čon* (c'wn)	*čan* (cnn)	*čiwant*	*čimant*	*čiwēδ*	*čimēδ*
δ-	*δon* (δ'wn)	*δan* (δnn)	*δiwant*	*δimant*	*δiwēδ*	*δimēδ*
ku-	*(ə)kū* ('kw)	—	*kiwant*	*kimant*	*kiwēδ*	*kimēδ*

The plural form of Er-deixis pronoun is also found: *pariwēšan*, etc. Those of Du-dexis are only sparsely attested with *ēδ*-extention: *čitēδ*, etc. Although fused with *δ-* 'with' *δiwant/δiwēδ* and *δimant/δimēδ* have genitive and locative functions. In Christian texts are attested *pariw* (pryw) 'on him', *čiw* (cyw) 'from him', and *δiw* (δyw) 'with him' of unclear origin.

Those which are construed with acc. case (of the light stem) are *par(u)*, *parm*, and *(ə)kū*, while *čan* (or *čon*) and *δan* (or *δon*) govern the inst.-abl. case.

3.1.4.2 Postpositions and circumpositions

The most common postpositions are *sā(r)* (s'r, C s') 'toward, from', *parēw* (pr'yw) 'together with', and *piδār*) 'for the sake of, because of' which themselves are often preceded by prepositions: *kū/čan . . . sār; čan . . . sār/piδār, δan . . . parēw; kū δēn sār* 'to the religion', *čan wənī sār* 'from him'.

3.1.5 Adverbs and interjections

3.1.5.1 Adverbs

Adverbs are indeclinables: *tīm* 'moreover', *yunēδ* 'immediately', *žγart* 'quickly'. Some adverbs take the optional ending *-ī* (< oblique ending): *rāmant(-ī)* 'always', *nūr(-ī)* 'today', *xwaδkār(-ī)* 'alone'. Oblique case forms of heavy stem nouns behave as adverbs or postpositions: *miδān-ī* (< *miδān* 'middle') 'among'. They also appear in predicative position, *a(m)sāk-ī* 'suitable' (C 's'q-y; *a(m)sāk* 'equipment').

Adjectives in the neuter accusative of light stems may function as adverbs, e.g. *šir-u* 'well', *žaγ-u* 'very' (< *žaγ-* 'severe'), *wisp-u* 'entirely'.

3.1.5.2 Interjections

In Sogdian the following interjections have been encountered: (a) ō ('w, M 'w·) 'O!', (b) *ai* (M 'yy, C 'y) 'hey!', (c) *nāy* (n'y) 'lo, see!'.

3.1.6 Numerals

3.1.6.1 Cardinal numbers

TABLE 5.12: CARDINAL NUMBERS

1–9	11–19	10–90	100+	1000+	
ēw C yō	yōnts	δas(a)	sat-	zār(u)	'1000'
əδū ~ (ə)δw(a)	δwāts	wīst	δwēsate	pančzār	'5000'
əθrē, šē		šēs	šēsat	βrēwar	'10000'
čatfār	čatfārats	čatfars			
panč	pančaθs	pančās			
uxušu	C xušarts	xušašti			
əβt(a)	C βtats	əβtāt	C βtasat		
əšt(a)	əštaθs	C štāt			
naw(a)	nōts	nawāt			

Units precede the decades: *əβt-wīst* '27', *nawa-nwāt* '99'. Numbers close to the decade may be expressed by subtraction: *ēw kanpī pančās* 'one less fifty' = '49'. 'Two' distinguishes gender and case, thus masc. *(ə)δwa*, fem. *(ə)δwe*, gen.-dat. *δiβnu*. In 2, 7–10, forms without *-a* originate from proclitics.

Some numerals show an old genitive ending *-nu*: *uxušu-nu, dwāts-nu*, etc.

3.1.6.2 Distributives, multiples, and fractions

The distributive marker is *-ki* or *-kankī*: *βrēwar-kī* 'by ten thousands'. Multiples are expressed with *yāwar* 'time' or *wāfar* 'so much': *uxušu yāwar uxušu* 'six times six', i.e. '36'. 'Half' is by *nēmē*: *əδū nēmē* 'two and half'.

3.1.6.3 Ordinal numbers

Ordinal numbers: *əftam- ~ əftamīk ~ əftamčīk* 'first', *δəβti ~ δəβtīk* 'second', *(ə)štīk ~ čaštīk* 'third'. The ordinals higher than 'fourth' are derived from the cardinals with the suffix *-am(i)* or with *-mīk*: *čatfār-am ~ čatfār-mīk* 'fourth' *panč-am ~ panč-mīk* 'fifth', *əšt-ami ~ əšt-mīk* 'eighth', C *štfārwīst-mīk* 'twenty-fourth', *zār-mīk* 'one-thousandth', etc.

3.2 Verbal morphology

3.2.1 Stem formation

3.2.1.1 Present and past stems

Verb forms are based on two basic stems, present and past (in the following indicated by present/past). Historically they derive from the Old Iranian present stems and perfect participles in *-ta*. Synchronically the two stems may differ remarkably: *kun- ~ wan-/ (ə)kt-* 'do, make' (< **krn(a)u-, krta-*). Other stems are suppletive, e.g. *wāβ/wayt-* 'say', *āβar/āyat* 'bring', *x-/(w)māt* 'be'. The productive past formant is *-āt* (occasionally shortened to *-at*) and some verbs have both forms, i.e. two past stems: *wayt-* and *wācāt*

corresponding to the present stem *wāc* 'allow, let, send' and *fašt-* (< OIr. **frašta-*) and *psāt* of *ps-* 'ask'.

3.2.1.2 Imperfective stems

Certain verbs have a distinct imperfect stem. These are characterized by the preservation, or analogical extension, of OIr. augment. Diachronically, these are those verbs where the augment was preserved between original directional prefix and present stem, and later fused with the prefixal vowel to *-ā-*, *-ī-*: *patγōš* 'hear' > *pat-ī-γōš* (< OIr. **pati-a-gauša-*), *framāy* 'order' > *fr-ā-māy* (< OIr. **fra-a-māya-*). However, verbs with **ham-* > *an/m-* have the prefix *m-*: *anxaz* 'rise' > *m-anxaz*, which was later extended to verbs with initial **ā-*: *āβar* 'bring' > *m-āβar*. Those verbs beginning with the Old Iranian preverb **abi-* and **us-* show imperfect stems with *-ā-* and *-ī-* respectively: *əβžay-* (< OIr. **abi-ǰawya-*) 'increase (vi.)' > *βāžay*, *sxwāy* (< OIr. **us-xwāhaya-*) 'take up') > *sīxwāy*. Other verbs, without an inherited prefix, such as *βar-* 'bear', show no alternation between present and imperfect stems.

The formation of the imperfect stem was later extended to stems without preverbs: *snāy* (< OIr. **snāya-*) 'wash' > *sīnāy*, *ufs-* (< OIr. **hufsa-*) 'fall asleep' > *wāfs*.

3.2.1.3 Derived voice stems

One finds numerous pairs of transitive-causative and intransitive-passive present stems which are inherited from OIr.: *xwēr* 'feed' vs. *xwar-* 'eat', *āβar* 'bring' vs. *āβir* 'be brought', *fnēš* 'deceive' vs. *fneš* 'be deceived', *sōč/suγt-* 'burn' vs. *suxs-* 'be burnt', *kun-* 'do, make' vs. *kir-* 'be done, be made'. These morphological causative and passive stems are largely obsolete with the possible exceptions of the so-called inchoative stems derived from the past stem with the suffix *-s* like *suxs-*: e.g. *γafs-* 'labour', *wiγrās* 'be awaken' (see Weber 1970: 127–228).

3.2.2 Nominal forms

3.2.2.1 Infinitives and verbal nouns

Sogdian possesses two kinds of infinitives, the past and the present, and verbal noun (see Yoshida 1979).

3.2.2.1a Past infinitive

This form consists of the past stem with the ending *-e* (*-y*, LS), or *-ø* (HS): *βaγt-e* (*βaxš-/βaγt-*) 'give', *γōβāt* (*γōβ/γōβāt*) 'praise'. On the syntax of the infinitive see section 5.3.5.

3.2.2.1b Present infinitives

While the formation and usage of the past infinitive remains constant in all types of Sogdian, the forms of the present infinitive vary considerably from text to text:

(a)   In Christian texts one often finds infinitives consisting of the preposition *par* and the present stem with an ending *-u* (LS) or without ending (HS), as in *par 0βar-u* 'to give', *par wēn* 'to see'.

(b) In Manichaean (and some Christian) texts the present stem with the ending *-i* (-y, LS) or without ending (HS) serve as infinitives, as in *βaxš-i* 'to give', *framāy* 'to order'.

(c) In Buddhist texts the stem enlarged by the suffix **-aka* is often attested, as in *wan-ē* (wn'y) 'to do' and *šmār-ē* (šm'r'y) 'to think'.

3.2.2.1c Verbal noun

Mainly in Christian, Manichaean and other late texts one finds a verbal noun derived from present stems with a suffix *-āmantē* (-'mnty): *patfas-āmantē* 'reading'. In Buddhist texts the present infinitive functions as a verbal noun.

3.2.2.2 Gerund and participles

3.2.2.2a Gerund

Gerund: Present stems with the suffix *-kya* (-ky', LS) or *-kī* (-ky, HS), serve as a kind of gerund meaning 'having ~ed, ~ing': *βar-kya* 'having brought, bringing', *wāβ-kī* 'having said, saying'.

3.2.2.2b Present participles

1   Present participle (a). The most productive participle in Sogdian is that derived with the suffix *-ēnē* (LS) or *-nē* (HS) from the present stem, as in *θβar-ēnē* 'giving, giver', and *pattāw-nē* 'enduring'. Heavy stems often take light stem ending as well, as *pār-ēnē* (M p'rynyy) 'nourishing, nourisher'. This participle is mainly employed as an attributive adjective or agent noun:

> *fritāt ke   wənī   sāt šire   əktyē   parwēžnē*
> love REL the.G-D all good.G-D action.G-D nourish.PART
> 'love which is the nourisher of all the good deeds'.

2   Present participle (b). Predicative and adverbial function is often played by the form which is derived by the *aka*-suffix, as in *nīδ-ē* (S nyδ'y) 'sitting', but *-ēk* or *-ēsk* (-yq, -ysq) in Christian texts. The Christian form always combines with the inflected forms of *(w)māt*, the preterite of *x-* 'be', and functions as a periphrastic expression denoting progressive past: *δārēk māt* 'he was having'.

3   Forms provided with the old ending of present participle enlarged with the *aka*-suffix, i.e. *-antē* are also met with. But they seem to be fossilized and are no different from adjectives: *əžuw-antē* 'living', *βōδ-antē* ~ *βōδ-antč* (fem.) 'fragrant', *wēn-antē* 'visible'.

3.2.2.2c Passive participles

1   Future passive participle. This participle is derived from the present stem by means of *-ī* or *-(i)čīk* (-y, -(y)cyk): *βīr-ī* (S βyry) 'to be obtained', *sumb-čīk* (M swmbcyk) 'to be bored'.

2   Past (passive) participle. This participle is derived from the past stem with the *aka*-suffix, masc. *-ē*, fem. marked by *-č*: masc. *parθaγt-ē*, fem. *parθaγtč* (S prδ'γt'k, prδ'γtch) 'stretched' (pres. *parθenč*). Forms derived from intransitive verbs have a perfect meaning while those from transitive verbs have a passive perfect meaning: *taγt-ē* '(having) entered' (*tēs/taγt-* 'enter'), *βīrt-ē* '(having been) obtained' (*βīr/βīrt* 'obtain').

3.2.3 System of tenses, aspects, and moods

The Sogdian tense and aspect system comprises present, imperfect, preterite, and perfect. There are six moods: indicative, subjunctive, injunctive, optative, imperative, and irrealis.

The perfect tense is formed periphrastically by the past participle and the auxiliary verb.

Middle endings are almost all replaced by the active, only a few of them having survived. However, 3s imperfect and optative middle endings are formally so salient that they serve as starting-points for the analogical development of new paradigms.

Passive is also formed by the past participle and the auxiliary verb. (On the traditional description of the Sogdian verbal system see Gharib 1965.)

3.2.3.1 Forms based on the present and imperfect stems

3.2.3.1a Major conjugations

The best attested personal endings of the above-mentioned moods are shown in the following tables. The two verbs selected to serve for the sample paradigms are *wan-* (wn-) 'do, make', imperfect stem *wan-* (LS), and *patɣōš-* (ptɣwš-) 'hear', imperfect stem *pat-ī-ɣōš-* (HS). As is the case with the nouns, heavy stem verbs frequently take light stem endings, and in fact, the heavy stem endings *-u* of 1s and *-i* of 2s imperfect are those of the light stems.

TABLE 5.13: LIGHT STEM CONJUGATION, *wan-* 'do'

	Present	Imperfect	Injunctive	Optative	Subjunctive	Imperative
1s	*wan-ām*	*wan-u*	*wan-u*	*wan-el-ē-m*	*wan-ān*	
2s	*wan-e*	*wan-i*	*wan-i*	*wan-el-ya*	*wan-a*	*wan-a*
3s	*wan-ti*	*wan-a*	*wan-a*	*wan-e*	*wan-ā-t*	
1p	*wan-ēm*	*wan-ēm*	—	*wan-ēm*	—	
2p	*wan-θal-ta*	*wan-θal-ta*	*wan-θal-ta*	*wan-ēθ*		*wan-θal-ta*
3p	*wan-and*	*wan-and*	*wan-and*	*wan-ēnd*		

TABLE 5.14: HEAVY STEM CONJUGATION, *patɣōš-* 'HEAR'

	Present	Imperfect	Injunctive	Optative	Subjunctive	Imperative
1s	*patɣōš-ām*	*patīɣōš-u*	*patɣōš-u*	*patɣōš-ēl-ēm*	*patɣōš-ān*	
2s	*patɣōš-e*	*patīɣōš-i*	*patɣōš*	*patɣōš-el-ya*	*patɣōš-a*	*patɣōš*
3s	*patɣōš-t*	*patīɣōš*	*patɣōš*	*patɣwš-e*	*patɣōš-āt*	
1p	*patɣōš-ēm*	*pat-ī-ɣōš-ēm*	—	*patɣōš-ēm*		
2p	*patɣōš-θ(a)l-ta*	*pat-ī-ɣōš-θ(a)l-ta*	*patɣōš-θ(a)l-ta*	*patɣōš-ēθ*		*patɣōš-θ(a)l-ta*
3p	*patɣōš-and*	*pat-ī-ɣōš-and*	*patɣōš-and*	*patɣōš-ēnd*		

The endings *-ām, -ān,* and *-āt* of the heavy stem are occasionally shortened to *-am,* etc. In the texts written in Sogdian script, one occasionally finds the ending (-'ty) (subj. 3s), which is a historical spelling for *-āt* (cf. OIr. *-āti ~ -āt*). The plural endings of the subjunctive and injunctive are not distinguishable; or rather, the subjunctive function is expressed by the forms of the injunctive.

The difference between 2 pl. endings *-0(a)* and *-ta* are largely chronological; while earlier texts written in Sogdian script tend to prefer *-0(a)* (-δ(')), most Christian texts in Syriac script show almost always *-ta* (-t'), invariable with both light and heavy stems and indifferent to moods. Some texts do distinguish between *-0(a)* and *-ta* in that the former is used in modal forms other than the indicatives (see Gershevitch 1954: § 737).

In addition to the forms listed above, the imperfect stem with optative endings, called the optative imperfect, serves as a durative or iterative past, e.g. *wāpat-e* 'it was falling' (< *ōpat* 'fall').

MIDDLE ENDINGS
Survivals of the middle endings are: 1p *-ēman* (-'ymn); 2s optative *-ēš* (-yš), which is the base of 2p opt. *-ēš0(a)* (-'yšδ('); 3s pres. *-te* (-ty); 3s opt. *-ēt* (-yt); 3s impf. *-t(a)* (-t(')). Only the 3s present form retains its original middle-passive force: *wēn-te* 'is seen' vs. *wēn-t* 'he sees'.

3.2.3.1b New minor and secondary conjugations

These are innovative and include the following:

1 IMPERFECT MIDDLE
This derived from the 3s impf. middle ending (e.g. *wan-ta* 'he did', *ās-t* 'he took'):

1s *-t-u*, 2s *-t-i*, 3s *-ta*, 3p *-t-ant*, e.g. *ās-t-u* 'I took', etc.

2 OPTATIVE MIDDLE
This is modeled on the 3s *-ēt* to which new modal endings are attached; attested are:

1s *-ēt-u*, 2s *-ēt-a*, 3s *-ēt-e*, 1p *-ēt-ēman*, 3p *-ēt-ēnt*, e.g. *βīr-ēt-u* 'I may obtain'. This category has hitherto been called 'precative'.

3 IRREALIS
This is marked by the suffix *-ōt* (LS) or *-t-* (HS) to which the new modal endings are attached:

1s *-(ō)t-u*, 2s *-(ō)t-a*, 3s *-(ō)t-e*, 2p *-(ō)t-ēšta*, 3p *-(ō)t-ēnt*, e.g. 2p *warn-ōt-ēšta* (C wrnwtyšt') 'you might believe'.

4 ĀZ-IMPERFECT
This new imperfect is a kind of durative past. It is formed by adding the imperfect endings to the 3s past form *āz* 'was':

1s *-āz-u*, 3s *-āz*, 1p *-āz-ēm*, 3p *-āz-ant* ~ *-āz-ēnt*, e.g. *δār-āz-ant* 'they were holding'. These forms are particularly numerous in Christian texts and only three verbs are encountered in other texts: *əskw-āz* 'remained', *šaw-āz* 'was going', and *āz* 'was'.

5 ĒK-(W)MĀT IMPERFECT
As mentioned above, this imperfect is only found in Christian texts. It consists of inflected forms of *(w)māt*, the preterite *x-* of 'be', and functions as a periphrastic expression denoting progressive past, e.g. *dār-ēk māt* 'he was holding'.

3.2.3.1c Progressive or durative particle: skun, archaic -əštan

Progressive force is provided by the particle *-skun* (-skwn) following the inflected form of verb: *šaw-ām-skun* 'I am going'. This particle is added to the present, imperfect, optative imperfect, and the *āz*-imperfect, but is incompatible with the subjunctive, injunctive, or impeertive. Its form varies from text to text: *-əskun > skun > skən > sk*, and *-kən > ku > k* after the ending *-t* or *-ant: wāß-ant-k* 'they are saying'.

The archaic alternative form *-əštan* ('štn) is attested in a few Buddhist texts (see Benveniste 1966).

3.2.3.1d Future and past prospective particle: kām

Future sense is added by the particle *-kām* (k'm), which is suffixed to present, and less frequently to the subjunctive, optative, injunctive, imperative, and the *āz*-imperfect, e.g. *βaxš-ām-kām* 'I shall give'. The form of this particle also varies from text to text: (k'm, k'n, C -q', -g'(n) after nasal consonants).

> *nē parēθān-kām   ... āpatē   nē kunām-kām*
> not sell.SUBJ.1s-FUT ... in.bondage not make.PRES.1s-FUT
> 'I shall not sell (her) nor put (her) in bondage'.

Of particular interest is the combination with *āz*-imperfect found in the manuscript C2, where the construction has past prospective meaning: e.g. *zwart-āz-kā* (zwrt'zq') 'was about to return'.

3.2.3.2 Copula and existential verb

Sogdian presents a great variety of the stems of the verb meaning 'be'. Some are inherited forms and the others, mostly 3s and 3p, are formed analogically on the basis of the former. The copula also serves as the ending in the intransitive preterite.

Indicative forms (mostly inherited from OIr.):
1s *im* ('ym), 2s *iš* ('yš), 3s *(ə)sti ~ əst* (('ʔ)sty) and *āz* '3s was' (''z; *āz*-imperfect); 1p *ēm* ('ym) or *ēman* ('ymn), 2p *-sθ(a)* (-sδ', as preterite ending) and *sta* (imperative), 3p *ant* ('nt).

Some examples of analogical formation:
3s *x-ači* (xcy), *i-či* (('ʔ)ycy), 2p *an-sθ(a)* ('nsδ(')), *išta* ('yšt' based on 2s ʔyš), 3p *x-ant* (xnt), *əst-ant* ('st'nt).

Optative and subjunctive forms are also secondary:
opt. 3s *əst-e* ('st'y), *x-e* (x'y), *y-e* (y'y), *e* (''y), 3p *əst-ēnt* ('st'ynt), *x-ēnt* (x'ynt), *ēnt* (''ynt); subj. 1s *x-ān* (x'n), 3s *əst-āt* ('st't), *x-āt* (x't), *y-āt* (y't), *āt* (''t).

In principle, the forms based on *əst-* denote existence while those based on *x-* and *y-* (reminiscent of Er-deixis vs. Ich-deixis) are mainly employed as a copula or auxiliary verb (see Weber 2000).

Notice that the article *(ə)xu* sometimes functions as a copula in texts written in Sogdian script:

> *xa āžunδuγta   əxu*
> 'the child daughter COPU = the child is a daughter'.

The preterite forms (see section 3.2.3.3a) are based on the stem *(w)māt-* 'was, were' and *(ə)kt-* 'became'.

3.2.3.3 Forms based on the past stem

The Sogdian has the preterite and potentialis (see section 3.2.3.3c) which are based on the past stem followed by auxiliary verbs. The preterite of intransitive verbs consists of the past stem and the verb 'be' (see above on the copula) while the transitive verb takes *δār-* 'have' as an auxiliary.

3.2.3.3a Preterite

The following are the preterite conjugations of *β(w)-ləkt-* (LS) 'become', *ēs/āyat* (HS) 'come', *kun- ~ wan-ləkt-* (LS) 'do, make', *wēn/wēt* (HS) 'see'.

TABLE 5.15: PRETERITE CONJUGATONS

	vi./pass. LS 'became'	vi. HS 'came'	vt. LS 'did, made'	vt. HS 'saw'
1s	*əkt-im*	*āyat-im*	*əkt-u δār-ām*	*wēt δār-ām*
2s	*əkt-iš*	*āyat-iš*	*əkt-u δār-e*	*wēt δār-e*
3s	*əkt-i*	*āyat*	*əkt-u δār-t*	*wēt δār-t*
1p	*əkt-ēm*	*āyat-ēm*	*əkt-u δār-ēm*	*wēt δār-ēm*
2p	*əkt-asθ*	*āyat-asθ*	*əkt-u δār-θ*	*wēt δār-θ*
3p	*əkt-ant*	*āyat-ant*	*əkt-u δār-ant*	*wēt δār-ant*

The 3s of heavy stems does not take the auxiliary verb, while light stems show the ending *-i* (-y). Occasionally one finds 3s feminine forms in *-a*, and neuter forms in *-u*: *nižt-a* (C nyžt') 'she went out', *xwart-u* (xwrtw) 'it was eaten'. Note that the transitive preterite of the light stems differs from the heavy in that the former takes the ending *-u* (-w) before the auxiliary.

The 2p ending is sometimes *-išta*: *kt-išta* (C qtyšt') 'you became' and *wēδār-išta* (C wyd'ryšt') 'you saw'.

In late texts one sometimes finds the present stem instead of the past before the auxiliary *δār*: *pēz-δārt* (C pyzd'rt) < *pēz/pišt-* 'hit' (cf. pyštw-δ'rt). In some late texts intransitive verbs are construed with *δār-*: *βāw-δārt* (C b'wd'rt) 'he approached'.

Another late feature is the fusion of the past stem and the auxiliary *δār-*: *kθāram* (kδ'rm, C qθ'rm) < *əktu-δāram*) 'I did', *framāδārt* (C frm'd'rt) < *framāt-δārt* (M frm't δ'rt) 'he ordered'.

MODAL FORMS

These are obtained by conjugating the auxiliary verb: subj. *əkt-āt* (’krt-ī) 'he may have become', *ōžγəst-δār-an* (M ’wjγyst-δ'rn) 'I may have settled'.

ELLIPSIS

When preterite verbs are coordinated, it is sometimes only the last verb which has an auxiliary verb, all the others consisting only of the past stems:

āraxsāt   *əti* ... *βərtu*   *əti* ... *patškwāt-δār-t*
'he resorted to and   brought ... and   entreated ...'.

3.2.3.3b Ergative construction

The earlier ergative construction is found in the Ancient Letters and some other texts written in Sogdian script. Characteristically, the agent is expressed by the gen.-dat. of the noun or enclitic pronoun, while the auxiliary verb agrees with the object/patient, and thus formally corresponds to the intransitive preterite. Compare the following phrases in standard preterite and in formulaic archaic ergative, which occur at different places in the same text, both meaning '(my) father banished me':

    standard (active) preterite:

      rti-mī əxu   əptri ..   yarmyān əktu-δārt
      and-me the.NOM father.NOM . . . punished make-PRET.3s

    ergative:

      əwən   əptre   . . . yarmyān əkt-im
      the.G-D father.G-D . . . punished become.PRET.1s

3.2.3.3c Potentialis system

The potentialis system expresses both possibility and anteriority of the action expressed by the verb (see Beveniste 1954). Intransitive (or passive) verbs take the auxiliary *β(w)-/əkt-* 'become' while the transitive *kun- ~ wan-/əkt-* 'do, make'. Light stems end with *-u* (-w) or *-a* (-') whereas heavy stems take no ending (on the origin of the potentialis see Sims-Williams 2007):

    (1) Possibility

      nistu β-ām (nystw β'm)   *tayta β-ant-sk* (C tyt' bntsq)
      'I can sit';   'they are able to go into';

      əktu wan-ān ('krtw wn'n)   *parāyat β-ām* (pr"yt β'm)
      'I should be able to do';   'I can arrive'.

    (2) Anteriority

Another function of these forms is to denote anteriority in the temporal clause introduced by the conjunction *čānō* 'when', and in that sense it may be called the pluperfect:

    rti čānō əxu suδāšn waytu-wan-a rti əxu barāman žyart   manxaz
    and when the S.   said-POT   and the brahmin quickly rose-IPF

    'when/after Sudhāshn had said (it) the brahmin got up quickly'.

3.2.3.4 Forms based on the past participle

3.2.3.4a Periphrastic perfect

The forms of the periphrastic perfect consists of the past participle (i.e. the past stem plus *aka*-suffix) and the existential verbs. With transitive verbs the auxiliary is sometimes *δār*, similar to the transitive preterite, but the construction is much less frequent than those with the existential verbs which have passive meaning, and one may reasonably argue that the Sogdian perfect is passive by nature. Participles agree with the subject (or with the object of transitive verbs) in number and occasionally in gender, so that morphologically

speaking the construction behaves like a nominal than a verbal formation. Modal forms are obtained by inflecting auxiliaries. Some examples are:

pšātēt δār-ēm (C pš'tyt d'rym, < pš'y)   'we have cast (them) away'.
γūtč x-ān (γwtch x'n, < γw-)   'I (fem.) should have been conducting sin'.
āžitēt wmāt-and (M ʾjytyytt wm'tnd, < ʾjy) 'they had been born'.

3.2.3.4b Periphrastic passive

1 Passive construction
A periphrasitic passive is formed with the auxiliary verb *β(w)-lǝkt-* 'become' and the past participle which often agrees with the subject in number and gender. The construction is also nominal rather than verbal. Some examples are:

βīrt-ē (masc. sg.) *βw-e* (βyrt'y βwy)   'it should be obtained'.
ǝktč-a (fem. sg.) *β-āt* ('krtch β't)   '(if) it should be done'.
sfrīt-ēt (pl.)   *wβ-and* (M sfrytyt wβ'nd) 'they will be created'.

2 Agents
In the periphrastic passive, an agent is frequently expressed by (a) a noun phrase preceded by the preposition *čan* and *par* (non-human agent). However, (b) nouns or pronouns in the genitive-dative or oblique case are also used:

(a) *čan:*

xa ǝžγma ... čan māxyān sār   ǝktč-a β-āt
the deceit ... from M.   from done.F become.SUBJ.3s
'the deceit would be done by Makhyan'.

(b) Oblique case:

βžīk rēž-ī   par0ūt-ē ... ǝkt-im
evil desire-OBL burnt   ... become-PRET.1s
'I was burnt by evil desire'.

3.2.3.5 Voice and causativity

3.2.3.5a Passive and middle

Apart from the periphrastic passive discussed above, passive or middle voice is expressed morphologically, though not productive. The old force of middle voice is perceivable in the 3s present with the ending *-te*, which has a passive sense when used with transitive verbs, e.g. *wēn-te* 'is seen' vs. *wēn-t* 'he sees'.

Sogdian has inherited passive forms. Inchoative forms similarly function sometimes as passives (for both see section 3.2.1.3 above):

(a) Old passive:

rti xa zāy   zpart kira
and the ground pure was.made (< OIr. **kṛya-*, cf. *kun-lǝkt-* 'do, make')
'the ground was made clean'.

(b) Inchoative:

rti   čiméδ   piδār   δū   waxs-ti
and from.this because.of *dhū* is.said (cf. *wāβ/wayt-* 'say').
'for this reason it is called *dhū*'.

3.2.3.5b Causativity

No special means for expressing causativity is found in Sogdian. Some inherited present stems have causative meaning (see section 3.2.1.3):

rti   ō   barāman xwart əti   čašant xwēr
and the.ACC Brahmin food and drink feed.IMPF.3s
'he made the Brahmin eat (= fed him with) food and drink'.

3.2.4 Negation

Sentences are negated by placing either *nē* (ny) or *nā* (n') before finite verbs; the both are masked by an ideogram L'. The preterite differs from the periphrastic perfect in that *nē* precedes the main verb in the former and the auxiliary in the latter: *nē āyātēm* 'We did not come' vs. *xrāmtē nē wmāt* 'he had not proceeded'. While *nē* negates proposition, *nā* is a prohibitive particle: *nē wāβ-ām kām* 'I shall not speak', *nē niyāt-δār-t* 'he did not take', *nā ps-a* 'don't ask'. Sometimes negation is reinforced by the indefinite pronoun or negative indefinite pronoun, e.g. *ēδč maryārt nē sumbt* 'he does not bore any pearls'. *nē* and *əsti* combine to give *nēst* (nyst) 'is not, there is not'.

The negation of the imperfect differs from this pattern (see Sims-Williams 1996b). It is formed not with the imperfect, but with either the present indicative or the injunctive, occasionally preceded by the enclitic element *-β(i)* (-β(y)):

positive:	negative:
rti δūrī zāy šaw-a (impf.)	*rti-βi nē δūrī zāy šaw-t* (pres.)
and far land went	and-ENCL not far land goes
'he went far'.	'he did not go far'.

3.2.5 Synopsis of the system of tense, aspect, and mood

The following table presents a tentative synopsis of the Sogdian verb system, summarizing the discussion above. The system distinguishes non-past and past, imperfective and perfective as well as indicative and non-indicative forms, showing an over all balance. Non-indicative forms are perfective in their default meaning. The imperfective forms of the non-past and the past forms are formed by adding *skun*, which is optional in the cases of the *āz*-imperfect and optative imperfect.

The *āz*-imperfect is largely confined to some Christian texts, except for 3s forms; it corresponds to the *ēk-(w)māt*-imperfect of other Christian texts.

The perfective use of the present tense also has future meaning, which can be reinforced by the particle *kām*. Note that the periphrastic perfect and passive are still nominal constructions and are not included in this synopsis.

TABLE 5.16: SOGDIAN SYSTEM OF TENSES AND MOODS

		Perfective		Imperfective
<ind.>	non-past	present present	future meaning + *kām*; subj. (+ *kām*)	pres. + *skun*
	past	imperfect preterite		imperfect + *skun* opt. impf. (+ *skun*)
			āz-impf. + *kām* (past prospective)	*āz*-impf. (+ *skun*) *ēk-(w)māt* impf.
	non-past		Past	
<non-ind.>	subjunctive (+ *kām*) optative (+ *kām*) injunctive (+ *kām*) irrealis imperative (+ *kām*)		preterite subj. preterite opt. pret. inj.	

4 SYNTAX AND USE OF FORMS

4.1 Use of cases

4.1.1 Light stems

The use of the six cases of light stems is comparable to those of the Old Iranian counter-parts and Sogdian follows the nominative-accusative pattern. Notice that the inst.-abl. case forms never occur independently, being always preceded by the prepositions *čan*, *čon* (cnn, c'wn) 'from' or *δan*, *δon* (δnn, δ'wn) 'with'. However, the original situation is obscured by several changes that Sogdian underwent; even in the most archaic texts the tendency can be observed to confuse genders and case endings and to generalize the masculine nominative form for all the other functions. Some examples follow:

(a) Genitive-Dative:

    xu zātē əwən   əptr-e   andēk   δāre
    the son the.G-D father-G-D custom have.OPT.3s
    'the son should possess his father's characteristics'.

(b) Nominative and Instrumental-ablative:

    xu xar-i   δan xar-a   parēw   ranft
    the donkey-NOM with donkey-I-A together fights
    'a donkey fights with (another) donkey'.

(c) Locative:

    rti mana əwī   xwārant əpkaš-ya tēs
    and my the.LOC right   side-LOC entered
    'it entered my right side'.

4.1.2 Heavy stems

Employment of the direct and oblique cases is complicated and the most consistent feature of the oblique case is the purely negative one. It never occurs in any syntactic context which would require the nominative of a light stem (where heavy stems properly have the direct case). This is partly due to the multiple origin of the ending -*ī*, which goes back to *-*ya* (i.e. the endings of loc. sg. m., gen.-dat., inst.-abl., and loc. f. sg.) (see Sims-Williams, 1982: 72–73).

The situation is ameliorated to some extent by the frequent employment of the articles, which may precede both light and heavy stems indiscriminately, and the fact that the preposition never collocates with the article is in consonance with this case-marking function of the article.

Another possible way of compensating for the poverty of Sogdian case-marking is the use of *xēp0* 'one's own' as a marker of the genitive:

> *xunē patkarē xēp0 wiδβāγ*
> that statue its.own explanation
> 'the explanation of that statue'.

4.1.3 Direct object and animacy hierarchy

1   The forms assumed by the direct object of transitive verbs may briefly be surveyed. Usually it is represented by the accusative or direct case:

> *ərti ō xšēšpat βaγ-u waδēδ nīšēδand*
> and the.ACC Splenditenes.DIR god-ACC there seated.3p
> 'there they seated the god Splenditenes'.

Sometimes, through syncretism discussed above a nominative form is substituted for the accusative:

> *xurn-i āse-skun*
> blood-NOM she was taking
> 'she was taking blood'.

2   Partitive sense is expressed by the preposition *čan* 'from':

> *nē čan yātē xwart-δāre*
> not from meat you ate
> 'you did not eat meat'.

3   As in other languages a verb meaning 'fear' takes the ablative object:

> *əžγu paČkwērt-skən čan xūβ*
> very much fears-DUR from lord
> 'he fears the lord very much'.

4   The two highest in the 'agency hierarchy', 'me' and 'you' (sg.), show special forms for the direct object *tāmā* and *tāfā*:

> *əzu tāfā wēnām-kām*
> I you.2s see. 1s shall
> 'I shall see you'.

5   Direct object marking

Regarding the animacy hierarchy, the latest stage of Sogdian, the language of the Christian Sogdian ms. C5, employs the oblique case for marking the direct objects which are both human and definite:

fšam-dārt   xēp0   zātē-ī
send.PRET.3s   his.own son-OBL
'he sent his son'.

Compare the ordinal direct case for the accusative function: *kū zātē sā* 'to the son'. One even finds this use applied to the oblique form of the 1s pronoun: *manā-y-ī* (C mn'yy) 'me (direct object)'. On the unique oblique case form which has developed in the language of C5 see section 7.2.1.2.

4.2 Use of verb forms

4.2.1 Possession

Possession is expressed by the gen.-dat. (or obl.) of the noun or (suffixed or independent) pronoun and the word for existence: (a), (b). Though not common, *δār* 'have, hold' is also known to serve as the verb: (c).

(a) *rti-βaγ   tawa   īnč   əsti əti rinčak zākt*
     and-lord!   you.OBL woman   is   and little children
     'and, O lord! you have a wife and little children'.

(b) *rti-mī   nēst   waδu nē zātē nē δuγta*
     and-me   is.not   wife not son not daughter
     'I have no wife, son, or daughter'.

(c) *rti   paδka xači əti   əxu zātē əwən   əptre   anδēk   δāre*
     and law   is   that.CONJ the son the.G-D father.G-D manner have.OPT.3s
     'it is customary that a son should possess his father's manner'.

4.2.2 Use of tense and aspect

The basic tense distinction in Sogdian is between the past and non-past. The former is represented by the preterite, imperfect, optative imperfect, *āz-* imperfect, and *ēk (w)māt* imperfect, the latter by the present tense. The present and past perfect forms describe the state reached through an action. The future is just a subcategory of the non-past tense and expressed by the present tense form occasionally followed by the particle *kām*. The future meaning is also conveyed by the subjunctive forms.

In Sogdian there is also the opposition between perfective and imperfective (or durative) aspect. In the non-past tense, the former is represented by the plain present form while the durative is marked by the particle *skun*. In the past tense, the imperfect and preterite represent the perfective aspect, while *āz-* imperfect, *ēk (w)māt* imperfect, imperfect followed by *skun* and optative imperfect, occasionally accompanied by *skun*, represent the imperfective (or durative) aspect.

It should be mentioned that the term 'imperfect' is used because the forms go back to the Old Iranian imperfect, and does not refer to the actual function. In the last stage of Sogdian, the form with *skun* became increasingly common, and virtually replaced the simple present form, similar to the development of Early New Persian *hamē* to

the imperfective marker *mī-*, which became obligatory in the indicative (see Telegdi 1938: 222–227). The bleaching of durative meaning is paralleled by the gradual reduction of form from *(ə)skun* to just *sk* or *k*.

4.2.2.1 Examples

(A) Plain form

1 Present:

> *ərti sāt   wispu ark   γərβ-ām*
> and whole all   work I. know
> 'I know every kind of work'.

2 Imperfect:

> *rti xu nawē nyiγōðan pat[ī]menč*
> and he new   garment he.wore
> 'he wore a new garment'.

3 Preterite:

> *əpyār   xwaβnu wēt-[ðār-]ām*
> last.night dream   see-PRET.1s
> 'last night I saw a dream'.

4 Optative as imperfect:

> *čan   əštana əxšiβti rōš[-e]*
> from breast milk   flowed
> 'milk was flowing from the breast'.

5 *āz-* imperfect:

> *par ēw žamnu əxu bagaβām əskw[-āz] əwī   rājgarī   mazēxak kanθī*
> in one time the Bhagavan stayed   the.LOC Rājagṛha great   city
> 'one time the Bhagavan was staying in the great city of Rājagṛha'.

6 *ēk (w)māt* imperfect:

> *yānt ēšt   žātēt   b[-ēk māt]-ant*
> these thing.PL said.PL being were
> 'these things were being said'.

(B) Forms with *skun*

1 Present:

> *rti əču wānō zārī   rāy-e[-əskun]*
> and why thus miserably you.cry-DUR
> 'why are you crying so sorrowfully?'

2 Imperfect:

> *ō   buti   t[ī]kōš-ant[-əskun]*
> the.ACC Buddha   they.saw-DUR
> 'they were looking at the Buddha'.

3   Optative as imperfect:

> *par patēδ   šaw[-ē]-nt[-əskun]*
> for begging they.went-DUR
> 'they were going for begging'.

4   *āz*-imperfect:

> *šir   andōxčnāk əti   šmārēkēn   šaw[-āz-skun]*
> very sorrowful and thoughtful was.going-DUR
> 'he was going in great anxiety and deep in thought'.

4.2.2.2 Imperfect and preterite

The difference between the imperfect and preterite is somewhat comparable with that of the German Präteritum and Perfektum (see Paul 1997). Thus, in one narrative text, the *Vessantara Jātaka*, the imperfect is used in the narrative part whereas the preterite is preferred in direct speech. The story begins with the following sentences:

> *māθ patīškway   wiyaša əti   βaya xwatāw pāruti   γrān   əktim*
> thus say.IMPF.3s rejoice! COMP lord! king!   for.COMP pregnant become.PRET.1s

> *paraw   māθ əti   əpyār   xwaβnu wēt-δārām*
> because thus COMP last.night dream   see-PRET.1s
> '(the queen) said (imperfect): "rejoice, O lord king!, for I have become (preterite) pregnant, because last night I saw (preterite) a dream" '.

In later texts the distinction became less and less clear, and one sometimes finds texts where imperfect and preterite forms alternate without any semantic distinction. For example, Christian text C5 prefers the preterite, and even in the narrative part one often finds the preterite forms:

> *ət   γōž-δārt   piδār ət   nīpēs . . .*
> and request.PRET.3s tablet and write.IMPF.3s . . .

> *ət   wiδās-δārant   wispu āδē*
> and be.surprised.PRET.3p all   someone
> 'and he requested (preterite) the writing table and wrote (imperfect) . . .
>   and all the people got surprised (preterite)'.

The merger of the two tenses may also be inferred from those preterite forms, though not common, which show the augment: *z-ī-wastant* 'they returned' (< *zwart/zwast-*) and *w-ā-māt* for *wmāt* 'was'.

To note, originally, the preterite tense functioned as a 'perfect', but when it lost its original function the perfect sense came to be supplied by the newly formed periphrastic perfect tense.

4.2.3 Use of modal forms

4.2.3.1 Functions

1   The imperative is used for command and prohibition, the latter being preceded by *nā*.

2   The optative mood has several functions:

(a)   In main clauses it denotes exhortation, comparable with the imperative, likewise with prohibitive *nā*.
(b)   It also expresses hypothetical sense, extended to parabolic use.
(c)   In subordinate clauses it mainly conveys hypothetical sense and is typically used in purpose, (indefinite) relative, and conditional clauses.

3   The subjunctive has the following functions:

(a)   It is used in both main clauses and in temporal clauses to express future.
(b)   It is used in purpose, (indefinite) relative, and conditional clauses for hypothetical action.
(c)   It also denotes injunction to the third person, and when negated *nā* is employed.

4   The usage of the injunctive mood has been studied by Sims-Williams whose findings are summarized as follows (Sims-Williams 1996b: 183):

(a)   It expresses the simple past in negative sentences.
(b)   It is used in conditional and purpose clauses.
(c)   It expresses politeness.

5   The irrealis mood denotes irreal or counterfactual situations in both main and subordinate clauses. Note that the forms identified by Gershevitch as 'irrealis ii' appear to be forms of *wmāt* followed by modal forms of the auxiliary: *wmāt-e* 'might have been' (wm't'y, opt. 3s) and *wmāt-u* 'I would have been' (wm't'w, inj. 1s).

4.2.3.2 Merger of modal categories: non-indicative

It is noteworthy that the subjunctive, optative, and injunctive are all used in subordinate clauses with verbs which also take the infinitive construction. This situation may suggest that the three categories were becoming confused and were merging into a single modal category, which may be referred to as non-indicative. For this tendency, note also the new, secondary modal endings, 1s *-u*, 2s *-a*, 3s *-e*, which are the endings of the injunctive, subjunctive, and optative, respectively.

4.2.3.3 Modal particles

4.2.3.3a Hypothetical particle xāt

xāt (x't), 3s subj. form of *x-* 'be', is sometimes added to inflected forms to give or reinforce hypothetical meaning: subj. *mir-ān-xāt* '(if) I should die', preterite *kθār-ant-xāt* 'if they have done'.

4.2.3.3b Hypothetical particle: -(a)n

In the texts written in Sogdian script one finds an enclitic element *-(a)n* which reinforces the hypothetical meaning of the sentence. The verbs are in the present indicative, subjunctive, optative or injunctive, that is, those denoting non-past events:

> *rti-n   pīšt kaδ uttekin wānō mān β-āt*
> and-N but if   U.   thus mind be.SUBJ.3s
>
> *əti   əma čatta waδu nē   δār-āt-kām   pāruti-š-an wāč-āt-kām . . .*
> CONJ the C   wife not have.SUB.3s-FUT but-her-N divorce.SUB.3s-FUT
> 'however, if Uttekin should think that he would not keep Chatta as his wife but would (rather) divorce her . . .'.

4.2.3.4 Examples

Here follow examples of modal forms (see also Complex sentences, section 5.3):

1   Imperative:

> *əti   xēp0 βrāt   par žəγma nā   ōzām*
> and own brother on lie   PROH condemn.IMPR.2s
> 'do not condemn your brother with lies'.

2   Optative:

> *rti-šu par šir   paδiβārčyā βarya*
> and-it on well enquiry   bring.OPT.2s
> 'You shall make extensive investigation on it'.
>
> *rti kaδ əxu wēšparkar xe   rti-šī   0rē rīt əskwe*
> and if he Weshparkar be.OPT.3s and-him three face exist.OPT.3sg
> 'if he should be Weshparkar, he should have three faces'.

3   Subjunctive:

> *rti əxu čan mana tans   mirant-kām rti mana mazēx əktānī βāt*
> and he from my affliction die.FUT.3p and me great sin   be.SUBJ.3s
> 'they will die because of affliction to me and I will have (committed) great sin'.
>
> *rti āδē . . . xwatkāmē nā   əskwāt*
> and someone . . . selfish PROH remain.SUBJ.3s
> 'One should not be selfish'.

4   Injunctive:

> *mana wānō wāβ čan tawa rēž   βēk-ti   ēč   nē raxnu   əkte*
> me   thus said from your desire outside-COMP something not venture.INJ.1s to.do
> 'he said thus to me: "I would not venture to perform anything except your wish".'

5   Irrealis:

> *rt-šu   tayu kaδāč   wā-xūpt nē frēštu-kunōte*
> and-him you not.at.all so-good not send-POT.IRR.2s
> 'you would never been able to send him such a good (letter)'.

6   Hypothetical particle *-(a)n*:

> *γarf   wātδār βōt-kām ke-t-(a)n   ... šawāt   rti-n ... yrān   arn   γawāt*
> many beings be.FUT   REL-COMP-*n* ... go.SUBJ and-*n* ... heavy sin   commit.SUBJ
> 'there will be many beings who would go ... and would commit great sins'.

4.2.4 Modal constructions

4.2.4.1 Necessity construction

Necessity is expressed by the auxiliary verb *sāč* 'be fitting' followed by the infinitive. The construction is impersonal and the agent, if expressed, is encoded by the enclitic pronouns or nouns in G-D or Obl.:

> *rti   mana sāčt   ite   čan   γara   kū γaru*
> and I.Obl. is.fitting to.go from mountain to   mountain
> 'I must go from mountain to mountain.'

4.2.4.2 Desiderative construction

Desire is expressed by the verb *kām* similarly followed by the infinitive. In Christian texts, *par* infinitives are used in a way similar to the past infinitives:

> *ke   kāmēnt   par-patxwāy wǝnī*
> who(REL) desire.OPT.IMPF.3p *par*-kill.INF   him
> 'who were wishing to kill him'.

4.3 Concord and agreement

4.3.1 Light and heavy stems

Naturally light stems follow the same agreement patterns as observed in Old Iranian. In contrast with the light stems, heavy and contracted stems are very relaxed with regards to the rules of agreement. To a large extent, the suffixes *-ī* (obl.) and *-t* (plur.) are treated as optional, being often omitted where clarity is unimpaired. When light and heavy stems are combined to constitute a phrase, each member behaves according to its own syntactic function, although the preference of nominative or direct case for every function is often encountered:

> *ō   tawa   roxšnu   čašmu*
> the.ACC   you.OBL light.ACC eye.ACC
> 'your light eyes'.

> *par   šukča   naβtča   zāy*
> on   dry.F.ACC   wet.F.ACC ground.F.DIR
> 'on the dry (and) wet ground'.

> *čan   tawa   zārčanūk   βaya*
> from you.OBL   merciful.DIR god.I-A
> 'from you, merciful God'.

əwī   sāt   tərxaki   tamya
the.LOC   all.DIR   harsh.NOM   hell.LOC
'in the whole harsh hell'.

4.3.2 Group inflexion and nouns after cardinal numbers

4.3.2.1 Group inflexion

One also finds the phenomenon of 'group inflexion' where only the last in a series of (usually asyndetically coordinated) words is inflected: *čan anxar paxarē-t-ī βēk* 'with the exception of fixed stars (instead of expected pl. obl. form *anxar-t-ī*) and planets'. The rule of group inflexion is not compulsory either. Thus, in combination with heavy stem adjectives and nouns one finds all the following examples in C2 (see Sims-Williams 1982: 69). The tendency is that determining adjectives are uninflected while descriptive adjectives agree in number:

βayānīk a(ṃ)βarz	'divine visitation'	(dir. + dir.)
γarf γampan-ī	'much labour'	(dir. + obl.)
sāt δēw-t	'all demons'	(dir. + dir. pl.)
xēp0 ančaman-t-ī	'one's church'	(dir. + obl. pl.)
a(ṃ)βasč-ī pačukān-ī	'constant prayer'	(obl. + obl.)
rāzyān-t wāxš-t	'symbolic words'	(dir. pl. + dir. pl.)
mazyātar-t ēš-t-ī	'greater things'	(dir. pl. + obl. pl.)
pēnamčīk-t-ī ēš-t-ī	'former things'	(obl. pl. + obl. pl.)

4.3.2.2 Nouns after cardinal numbers

Both singular and plural forms follow cardinal numbers, while light stem nouns sometimes take numerative forms in that position:

čatfār δβar-a (num.) 'four doors'   vs.   *δwāts δβar-ta* (pl.) 'twelve doors';
əβt paxarē (sg.) 'seven planets'   vs.   *əβt paxarēt* (pl.) 'id'.

4.3.3 Agreement of the subject and the predicate

The number of the predicates usually agrees with that of the subject. However, 3s, which is the unmarked number and person, occasionally appears instead of the expected plural form: *əžyantt xartēt əsti* 'the envoys have gone (sg. *əsti* instead of *əstant* or *xant*)'.

The relaxed nature of agreement in Sogdian verbs may also be exemplified by the 3s fem. forms of the intransitive preterite of light stem verbs. One sometimes finds special feminine forms ending with *-a* (-'). But ordinary forms without gender distinction are commonly encountered:

xu   nafša   pyamt-a   əti   əkt-a   anγatē   əpu   xwēč
the   Nafsha   was.healed.F   and   became.F   whole   without pain
'Nafsha was healed and became whole and without pain'.

rti   īnč   βēksār   nižt-i
and   woman   outside   went.out
'the woman went out'.

5 SYNTAX OF SIMPLE AND COMPLEX SENTENCES

The syntax is the least studied field in the study of Sogdian grammar. As yet there exists no general account of Sogdian syntax. On some topics see Heston 1976.

5.1 Order of constituents

5.1.1 Noun phrase structure

In Sogdian the head noun of noun phrases is preceded by the article, demonstrative, possessor noun, and adjective:

(a) *ō tava roxšnu čašmu*   'the your light eye'   = 'your light eye'
(b) *yunē mana xēpθ žyāwar*   'this my own heart'   = 'this heart of mine'
(c) *par xēpθāwandī framān*   'on owner's order'   = 'according to the owner's order'
(d) *xu xunax γrīw*   'the that body'   = 'that body'.

Similarly, cardinal numbers and words for unit precede nouns with this order:

   nōts kapčakk nā-art yantam
   '19 kapchak of non-ground wheat'.

Nouns designating rank, title, or category usually follow the proper nouns: *smēr yari* 'Mount Sumeru', *uparatt šamani* 'Monk Uparatt'. However, counter examples do exist: *šamani yansyan* 'Monk Yansyan', *wispišē suδāšn* 'Prince Sudhāshn'.

A prepositional phrase modifying a noun follows it:

   ō   zernēnē wartan δan anyatč parštāk
   'the golden chariot with all the equipment'.

When an appositional phrase is governed by a preposition, the latter is repeated before each member (see Sims-Williams 1973):

   čan   širβaxč   əti əstapsarē čan   farnxunt   zātēt
   lit. 'from Shirvakhch and Əstapsarē from Farnkhunt's (two) sons'

5.1.2 Clause structure and word order

5.1.2.1 Complementizer

A syntactic feature peculiar to Sogdian (and for that matter to Bactrian as well) is that each clause (both main and subordinate) contains an enclitic complementizer (COMP): *(ə)t(i)*, *-uti* (= 'PZY, ZY). This stands in the second position from the beginning, to which other enclitic elements of the sentence are added (see Sims-Williams 1985b, Yakubovitch 2005). Complementation includes the most frequent *(ə)rti*, which marks the beginning of a clause and etymologically consists of the adverb *r-* (see Khotanese *rro* 'also') and *-ti*.

The same is found in relative clauses, where relative pronouns or adverbs are followed by the complementizer. Similarly, direct quotations are also treated as independent clauses, where the second position is occupied by the complementizer (see Weber 1971):

rti-šī   *xā xwatēn māθ patīškway māθ əti*   *βaya mēnu*
and.COMP-him the queen thus said   thus COMP lord! I.thought

čan xwēr-βayī əti   *əβt*   *čintāman ratni nīži*
from sun-god COMP seven cintāmaṇi jewel it.went.out
'the queen said to him: "O Lord! I thought thus: <from the sun god went out the seven cintāmaṇi jewels>" '

As shown by *čan xwēr-βayī əti* above, the element standing before the complementizer is not always a single word but one syntactic unit, whereby the vocative form is not counted.

However, this feature was declining. Thus, in such a late text as Christian manuscript C5 the complementizer is almost lost, the only vestiges being the -*t* element found in the compound conjunctions *kat* (C qt) 'that, who (rel.)' and *pāt* (C p't) 'for (conj.)', and interrogative particle *čut(i)* 'est-ce-que'.

5.1.2.2 Enclitics

Even as a Middle Iranian language. Sogdian still observes the Wackernagel's law, according to which enclitics occupy the second position in the sentence. As stated above, in classical Sogdian the second position is usually occupied by the complementizer *(ə)t(i)* and other enclitics are added to it:

pār-ti-šī   *xu wīnā žani*   *framātδāre*
but/for-COMP-him the lute to.play you.ordered
'but you ordered him to play the lute'.

When in a very few cases one pronominal suffix is attached to another, the first person precedes the second and the third: *ərti-m-f-(a)č* ('ḤRZYmβc) 'to me from you', *rti-m-šu* ('ḤRZYmšw) 'and him for me'.

Apart from the enclitic pronoun, several other enclitic elements occur in Sogdian. They are prepositional elements (-c, -δ, -t), -*βay* 'O lord!', and -*n* (hypothetical particle), -*β(i)* (perfective particle employed in the negated imperfect sentence), etc. These enclitics follow pronouns:

māθ əti-š-n   *uβyu šamani Yansyan xwati*   *əti . . . par kāmē xawāt*
thus COMP-her-PART both monk Yansyan himself and . . . by desire hit.SUBJ.3s
'so that either monk Yansyan himself or . . . may hit her as he likes'.

5.1.2.3 Word order

5.1.2.3a SOV

Sogdian is basically an OV language where heads follow the dependent elements. Thus, the basic structure of the Sogdian sentences consists of SOV. However, Sogdian is far from being a consistent OV language and attests a number of counter-examples. The unmarked order is (a) SOV, but (b) the verb initial order is also often encountered. Extended constituents, including subjects and objects, tend to follow shorter constituents: (c), (d). Sometimes, even a part of a longer constituent is extraposed toward the end of a sentence: (e). Standards of comparison precede the adjective: (f). Adverbial elements tend to stand before verbs or objects of verbs: (g) On the orders of the adposition and the auxiliary verbs see sections 3.1.4 for adpositions and 3.2.3.3–4 for auxiliaries.

(a) *ərti xu maryārtī xēpθāwand . . . ō   satu δīnār zern tōž*
and the pearl's   owner   . . . the 100   denar gold paid
'and the owner of the pearls paid the hundred gold denars'.

(b) *rti   frāyāz əxu wispišē Suδāšn   wispu āδč βayte*
and began the prince Sudāshn all things to.give
'Prince Sudāshn began to give all the things'.

(c) *rti-šī   ēs   čan patīčsār ēw brāman čanstī šāw əksi əti . . .*
and-him came from in.front one brahmin very   black lean and . . .
'one very black, lean, . . . brahmin came towards him'.

(d) *rti   əxu wēn yarf bikšē   nāf   əke-ti   par patēδ   šawēnt-əskun*
and he   saw many mendicant people who-COMP for begging were.going-DUR
'he saw many mendicant people who were going for begging (alms)'.

(e) *ō   nāft   martsār tūx   əfšām   əti arkkarēt*
the people.PL hither   quickly send.IMPV and workers
'send people and workers hither quickly'.

(f) *čan   šakara nāmartar*
from sugar sweeter
'sweeter than sugar'.

(g) *rti yonēδ   βōk   βīrant-kām*
and immediately salvation they.find-FUT
'and they will get salvation immediately'.

5.1.2.3b Counter-examples

Since Sogdian is not a consistent OV language, counter-examples are not uncommon. But some of them are due to the prototypes. The bulk of the Sogdian texts are translations from other languages, and the word order often reflects those of the original language or texts. This is particularly evident in translations from the Syriac New Testament. See the following example. Notice, however, that the deep-rooted order of possessor + possessed (*βar-ēt-ī žamnu* 'fruits' time') is never inverted:

Syriac
kd   dyn mt'   [zbn'd- p'r] šdr 1-['bdw-hy]   1wt   plḥ'
when but reached [time of-fruit] sent to-[servants.his] towards farmers

Sogdian
pīšt čānō parāyat [βar-ēt-ī žammu] fšamδārt [xēpθ βantētī] kū βāy-δārēt
but when reached [fruits' time]   sent   [his   servants] to farm-keepers
'however, when the time of fruits (i.e. harvest) came he sent his servants to the farmers'.

Naturally, emphasized and topicalized elements tend to be placed towards the beginning of sentences:

rti xa uxušu βrēwar īnč   mart mart čan əštana əxšiβti rōše
and the six   10,000 woman each each from breast milk   flow.OPT.IMPF.3s
'as for 60000 women, milk was flowing from each woman's breast'.

5.2 Questions

Sogdian does not seem to have a special device for forming interrogative sentences. Interrogative pronouns and adverbs follow the ordinary word order: (a), (b). However, since

the interrogatives are the topics of the statements in their very nature they tend to be placed at the beginning of sentences: (c). Yes/No questions are formed either by an ordinary order (possibly with different intonation): (d), or by putting *čut(i)* at the beginning of a sentence, of which -t(y) originates from the complementizer. Disjunctive sense is sometimes emphasized by the phrase *katār (əti) nē* 'or not?' placed after the sentence: (e). Rhetorical questions make use of the particle *p(u)nūkar*: (f).

(a)   *taγu pērnamstar ču   əktya k0āre*
       you before   what deed do.PRET.2s.
       'What deed did you do before?'

(b)   *xa əspiyi-ti   kutsār wāčām*
       the horses-COMP   where I.send
       'Where shall I send the horses?'

(c)   *ərt-šu   čakanāč piδār   marγārt nē   framāye suβte*
       and-him   whence because pearl   not order.2s bore.INF
       'Why didn't you order him to bore pearls?'

(d)   *nē   tkōš*
       not see.INJ.2s
       'Don't you see?'

(e)   *čuti   xa zākt   taδēδ āyatant katār-əti nē*
       whether the children there came   or-COMP not
       'Have the children arrived to you or not?'

(f)   *əčuti   pnūkar tawa   wānō nē   patγōšti*
       whether pnwkr by.you thus   not heard.PRET
       'Have you never heard thus? (Yes, you have surely heard thus.)'

5.3 Complex sentences

In Sogdian there are three kinds of subordinate clauses: (a) relative, (b) adverbial, and (c) nominal.

5.3.1 Conjunctions

Coordinate conjunctions are *(ə)rti* (('')rty, 'ḤRZY) which marks the beginning of the clauses, *ət(i)* ('t(y) = ('PZY, ZY) 'and', and *katār (əti)* (kt'r, 'WZY) 'or'. Apart from these, several adverbs placed towards the sentence initial are employed to enhance cohesion. Some of the most frequent are: *mas* 'also, again', *pīšt* 'but', *nūkar, patsār* 'then', *ūm* 'furthermore'. Of Old Iranian origin are connecting particles *βāδ . . . βāδ* 'sometimes . . . sometimes', *uβγu . . . uβγu* (wβyw, C by) 'both . . . and . . .'.

  Subordinate conjunctions are:

1   preceding the main clauses: *ču* 'if', *čānō* 'when, while, since', *kaδ(a)* 'if', *mant* 'when, while';

2   following the main clauses (*əti* is the complementizer described above): *ət(i)* ('t(y) = 'PZY, ZY) 'that', *čānō ət(i)* 'as (= like), than', *pār(u)ti* (p'r(w)ty, later *pāt*) 'for, (not . . .) but', *parəw ət(i)* (pr'w 't(y)) 'because', and *yiwār-əti* (yiw'r ('ty)) 'however'. *kaδ* and *ət(i)* on the one hand, and *māθ* (m'δ = KZNH) 'thus' and *ət(i)* on the other, are combined to give the new subordinate conjunctions *kat* (kt, C qt) 'that' and *māt* (m't) 'that', respectively, which are quite common in Manichaean and Christian

texts. These conjunctions occasionally combine with adverbs to form compound conjunctions as *čiwēδ piδār pāruti* 'because', lit. 'because of that for', *paraw mā0 əti* 'because', *witūr əti* (wytwr 'ty) 'until', etc.

5.3.2 Relative clauses

5.3.2.1 Correlatives and gaps

Relative clauses are introduced by relative pronouns or adverbs, to which is attached the complementizer *ət(i)* in Classical Sogdian. They always follow their head, which is usually marked by a cataphoric demonstrative. Although the following remark is not without exception, a general tendency is observable: the correlative in the relative clause is the verbal ending if nominative (a), while the direct object is gapped (b). With other cases, the relative may be in the oblique itself *(ə)kya* (c), but generally it assumes the relative particle *(ə)ke*, while the correlative is a pronoun or adverbial demonstrative filling the gap in the relative clause (d), (e) (see Provasi 1997). Comparative clauses can be classified as relative clauses (f):

(a)   *ōnō martī wiru   kunāt   ke-ti-šī   xwati rēžāt*
      that man husband make.SUBJ.3s REL-COMP-herherself please.SUBJ.3s
      'she shall make that man her husband who might be pleasing to her'.

(b)   *yunē čakraβart   čintāmani dārani ke-ti   əzu parβērāt-δārām*
      this Cakravarti Cintāmaṇi spell REL-COMP I   explain-PRET.1s
      'this Chakravart Chintamani spell which I explained'.

(c)   *əwən šē   ratne . . . əkya-ət-mī   sāče   xu pāš   pāt*
      the   three jewel . . . REL-COMP-me be.fitting.OPT.3s the honour observe.INF
      'three jewels . . . to which it may be fitting for me to pay honour'.

(d)   *xānd āfrītēt   ōtākt   ke əti-šan   sāk   əti patšmār nēst*
      those blessed places REL COMP-their number and number is.not
      'those blessed places whose number and counting do not exist'.

(e)   *xānd āfrītēt   ōtākt . . .   kū-əti   waδēδ mēnand xa roxšnda βayīšt*
      those blessed places . . . where(REL)-COMP there stay.3p the light   gods
      'those blessed places where the light gods are staying'.

(f)   *frətar māx ō   əβiža βarēm čānō-ti   əwən əptre   xā punyān zarnē*
      more we   the evil   endure than-COMP the   father's the merit lost
                                                                   βāt
                                                                   become.SUBJ.3s
      'we would rather endure sufferings than our father's merit should be lost'.

5.3.2.2 Other relatives

(ə)ke is originally a nominative singular form used with an animate noun, but it tends to be generalized to all the contexts and behaves like a relative particle. On this point see the examples cited in the previous section.

   In principle any interrogative pronoun or adverb can be used as a relative. Here follow several examples:

(a)   *wispu ark   ču əti-mī   tayu framāye*
      all   work which (REL) COMP-me you order
      'all the work which you order me'.

(b) *wāγōnē sāmār βīrt-δārant*
such samādhi obtain-PRET.3p
    čakanāč əti   ənyu βrēwar sāmār   βīrt-βāt
    from.which(REL) COMP other 10,000 samādhi obtain-PASS.POT.3s
    'they obtained such meditation from which other ten thousand
    meditations can be obtained'.

(c) *tīm āδču   əsti katār-əti-əmī   βaxše-kām*
still something is which(REL)-COMP-me give.OPT.2s-FUT
'still there is something which you could give me'.

5.3.2.3 Indefinite relative clauses

Indefinite relative clauses have no head and precede the main clause, which is often introduced by the conjunction *rti*: (a). Sometimes indefiniteness is reinforced by *xatu* and/or *āδparm*: (b). Another way of reinforcing indefinite force is to duplicate the relatives, first by the original (in the present case *kū* 'where') and again by *ke*, which in this function is a relative particle rather than a relative pronoun: (c). *ču žamnya*, lit. 'whatever time' functions as a kind of compound conjunction meaning 'whenever': (d).

(a) *rt-mas   katām   āžunī āžit   rti xu rāmant jātismar γrīw βīrt*
and-again which(REL) life   is.born and he always jātismara body gets
'again in whichever life he may be born, he will always get jātismara-body
(i.e. he will remember his previous lives)'.

(b) *əču   xatu āδparm wanti rti-šī   wisp sāt . . .*
what(REL) ever ever   does and-him every all . . .
'whatsoever he may do, everything is . . . for him'.

(c) *kū   sār   ke   šawēte   xu fuxār . . .*
where(REL) toward REL go.OPT.MID.3s the blessed.one
'wherever the blessed one may go . . .'.

(d) *rti ču   žamnya čan βaγistān ōxazāt   rti paru butyāk*
and what time.LOC from heaven descend.SUBJ.3s and on   buddhahood
                                                rā0 ōštayt
                                                road stands
'whenever he may descend from the heaven he will stand on the road of
Buddhahood'.

5.3.3 Adverbial clauses

Clauses that explicitly or implicitly state logical and temporal conditions precede the main clause, and are typically introduced by conjunctions such as *kaδ(a)* 'if', *čānō* 'when, since', *mant* 'when, while', *kū parm* 'as long as', etc. In classical Sogdian both subordinate and main clauses are introduced by a conjunction *rti*.

rt-šu   kaδ nē   waγtu-δāre rt-šu   yōnēδ   nā wāč
and-him if   not send-PRET.2s and-him immediately not send.IMPV.2s
'if you have not sent him, do not send him immediately'.

rt-šu   kū parm nē tōžān   rt-šu . . . paru wartu   δārām-kām
and-it as.long.as not pay.SUBJ.1s and-it . . . on   interest have.1s-FUT
'until I pay it (= as long as I do not pay it), I shall owe (lit. 'have') it with interest'.

Other adverbial clauses expressing purpose, cause, etc. are accompanied by appropriate conjunctions and follow matrix sentences. Here follow examples of (a) cause and (b) purpose:

(a) *čan   šē satu əfsānx zāy   əti   āyatim   paraw   māθ-əti   wānō patīyōš*
from 3 100 mile land COMP I.came because-COMP thus I.heard
'I have come from the place 300 miles far (from here), because I heard thus . . .'.

(b) *šawām-əštan əti-š(a)č*   *θβār   βīrān*
I.go-DUR   CONJ(=COMP)-him-from gift   obtain.SUBJ.1s
'I am going (there) so that I might get a gift from him'.

5.3.4 Nominal clauses and direct quotation

5.3.4.1 Nominal clauses

The most common nominal clause is the complement of a transitive verb which is introduced by *ət(i)*, *kat* (< *kaδ-ət(i)*), or *māt* (< *māθ ət(i)*): (a). An appositional clause is also introduced by the same conjunction: (b).

(a) *xa pīδt   wānō wēnant əti   xa brāmant   ō   . . . xwatāw   parēpant*
the elephants thus saw   that the brahmins the . . . king   took-away
'the elephants saw that the brahmins took the king away'.

(b) *rti   wānō paδka   nēst   əti   əwī   yantāk   rāθī   anpatē*
and thus judgement is.not that the bad   way fall.OPT.3s
'there is no such judgement that he should fall into a bad existence'.

5.3.4.2 Direct quotations

Direct quotations are treated differently from clausal complements in classical texts but later they came to take the same construction. Thus in one Buddhist Sogdian text *Vessantara Jātaka* direct quotations are always introduced by the complementizer placed in the second place (see section 5.1.2.1), while in later texts they are often introduced by *kat*:

ərt-xu   xēpθāwand wānō wāβ kat-βay   munō martī. . .
and-the owner   thus said that-lord! this man . . .
'the owner said: "O Lord! (I hired) this man . . ." '

Notice that in Sogdian it is a rule that the verbs introducing direct speech or nominal complements should be preceded by a pleonastic adverb *wānō* or *māθ* 'thus'.

5.3.5 Use of infinitives and verbal nouns

5.3.5.1 Past infinitives

Functionally speaking, infinitives and verbal nouns are comparable to subordinate sentences and their usage is described in this section.

The past infinitives are dependent on verbs such as meaning 'order', 'begin', 'wish', 'be fitting', etc. and always follow them. Below is the past infinitive dependent on an impersonal verb *sāšt/sāčt* 'it is fitting': (a). A subordinate clause dependent on *sāšt* is also given: (b)

(a) *rti əxu nē βaxše ōnō əču-ti-šī nē sāčt βayte*
and he not give.OPT.3s which (REL)-COMP-him not is.fitting give.INF
'he may not give that which is not fitting for him to give away'.

(b) *wānō sāšt kat par wēšant čūpar yān xōžēm čan βaya*
thus is.fitting that on them upon favour request.OPT.1p from god
'it is fitting that we should ask favour from the God upon them'.

One special usage of the past infinitive is to appear in the construction *čan . . . kārī* (cnn . . . k'ry) meaning 'after having . . .', where the light stem takes the ending *-a* or *-u* while the heavy is without ending (on the origin of the construction, see Sims-Williams, 2007):

čan murtu kārī   *mana čan nirβān witart kārī*
from dying after   my from nirvāṇa passing after
'after having died';   'after my passing into the nirvāṇa'.

5.3.5.2 Present infinitives and verbal nouns

As stated above the form and usage of the present infinitives in Sogdian vary considerably from text to text. Their unstable nature owes much to the fact that they were developing from the nominal form based on the present stem, i.e. productive verbal nouns, to the forms more similar to the past infinitives. *par* infinitives in Christian texts are used in a way similar to the past infinitives:

ke kāmēnt par-patxwāy wənī
who(REL)ʹ desire.OPT.IMPF.3p *par*-kill him
'who were wishing to kill him'.

The present infinitives of the Buddhist texts show more characteristics of the verbal nouns than *par* infinitives of the Christian texts, where the ordinary verbal nouns are derived by means of the highly productive suffix *-āmantē*.
Compare the Buddhist Sogdian and Christian functionally equivalent:

zanē žamnya   *wya zanāmantē mēθ*
giving.birth time.LOC   her giving.birth day
'at the time of giving birth';   'the day of her giving birth'.

Forms derived with *-āmantē* are very few in Buddhist texts whereas they are much more numerous in Manichaean texts.

6 LEXIS

6.1 Sogdian lexicon

The Sogdian lexicon consists of three groups of words: (a) those inherited from Old Iranian; (b) loanwords; (c) foreign elements temporarily appearing in texts. Some of the native Iranian elements show distinctively East Iranian features. Thus, *kaṃθ* 'town', *kap-* 'fish' are typically East Iranian. *farn* 'glory, fortune' has recently been shown to be originating from the language of Scythians at the Old Iranian stage (see Lubotsky 2002).

Foreign elements are mainly Sanskrit forms in Buddhist texts and Syriac words in Christian texts (see Sims-Williams 1983b, 1988). They are quite numerous and their

number seems to be dependent on the scholarship of each translator. On the other loanwords originating from the language contact, see section 7.3.

6.2 Derivation

6.2.1 Suffixes

Derivation in Sogdian, which is an OV language, is mostly by suffix. The most productive suffixes are the following. They are so productive that loanwords come to be provided with them:

(a) Adjectives from (mainly) nouns:

    1   *-ēnē*, f. *-ēnč* (-'yn'k; *aka*-stem) 'made out of ~': *zernēnē* < *zern* 'gold', *δārukēnč* < *δāruk* 'wood'.

    2   *-čīk, -īčk* (-cyk, -yck): *γarčīk* 'mountain-like' < *γar-* 'mountain'; *cāδ arčīk* 'low' < *čāδar* 'below'; *βēkpārčīk ~ βēkpārīčk* 'external' < *βēkpār* 'outside'. Compare the two adjectives derived from *zāwar* 'power': *zāwarčīk* 'helpful' vs. *zāwarkēn* 'powerful'.

    3   *-mēnč* (-mync): *īnčmēnč* 'female' < *īnč* 'woman', *žūkyāmēnč* 'safe, sound' < *žūkyā* 'health' < *žūk* 'healthy'.

    4   *-īk* (-yk): *δarmīk* 'of dharma' < *δarm* (< Skt. *dharma* 'Buddhist law').

(b) Abstract nouns:

    5   *-yāk* (-y'k, from LS), *-yā* (-y'kh, M -y', from HS): *roxšnyāk* 'splendour' < *roxšn-* 'light', *βayyākyā* 'divinity' < *βayyāk* 'id.'. < *βay-* 'god'.

    6   *-āwē* (-'w'k, *aka*-stem): *friyāwē* 'love' < *fri-* 'dear', *patpatēnāwē* 'isolation' < *patpatēn* 'opposed'.

    7   *-ōnī* (-wny): *mastōnī* 'drunkenness' < *mast* 'drunken', *tāyōnī* 'theft' < *tāy* 'thief'.

(c) Others:

    8   *-ānč* (S -'nch): female counterpart of nouns meaning male human: *upāsānč* < *upāsē* 'lay-brother' (< Skt. *upāsaka*), *nəγōšākānč* < *nəγōšāk* 'Manichaean auditor' (< Part. ngwš'g). For *upāsānč* see Tocharian B *upāsakāñca* 'id.'.

    9   *-āu* (-'w) 'in the ~ language': *čīnāu* 'in Chinese' < *čīn* 'China', *syuδyāu* 'in Sogdian' < *suγδ-* 'Sogd'.

6.2.2 Prefixes

Prefixes are much less common than suffixes. Productive are the following:

1   *āu-* ("w-) 'co-': *āupatyāp* 'sharer' < *patyāp* 'share', *āuxānēt* 'co-inhbitants' < *xānā* 'house'.

2   *mant-* (mnt-) privative: *mantγərβāk* 'foolish' < *γərβāk* 'wise'.

3   *nā-* (n'-) privative: *nā-paδkčīk* 'unlawful', *nā-martəxmē* 'non-human'. The past participle is negated with this prefix: *nā-wēt* 'invisible' < *wēn/wēt* 'see'.

4   *(ə)pu-* ((')pw-) privative: *pu-patšmār* 'innumerable' < *patšmār* 'number'.

5   *frī-* (S pry-) 'philo-': *frī-rwān* 'who loves one's own soul' < *rwān* 'soul', *frī-āβraxsē* 'lascivious' < *āβraxsē* 'lust'.

6.3 Composition

Several patterns of both verbal and nominal compounds are known (see Gershevitch 1945).

(a)   Containing present stems:

1   noun + present stem + *aka*-suffix forming agent nouns: *framān-patγōšē* 'servant, obedient' (*framān* 'order' + *patγōš* 'hear'), *yipāk-βarē* 'angry' (*yipāk* 'anger' + *βar-*'bear'). Notice that older type of the similar compounds show as its second member a slightly different form than the present stem: *əptxwār* 'carnivorous' < **pitu* + *xwāra*, compare *xwar-* 'eat', *širxōzē* 'friend' < *šir-* 'good' + *xwēž* 'to wish'. The following compounds contain Old Iranian acc. sg. forms as the first members: *βžangārē* 'sinful', *širankārē* 'pious'.

(b)   Containing past stems:

2   past stem + noun forming bahuvrīhis: *sußt-γōš* 'with pierced ears' < *sumb*/*sußt*-'pierce', *əkt-əspās* 'obedient', lit. 'made-service', *βart-γamband* 'enduring the strain (*γamban*) (pl.)' < *βar-*/*βart* 'bear'.

3   noun/adverb + past participle: *wāt-nyātē* 'wind-sick, taken by wind (*wāt*)' < *nyās*/*nyāt* 'take', *čon-žmē-miδānī-āyatē* 'coming from among (*miδān*) the anger (*žmē*)' < *ēs*/*āyat* 'come'.

(c)   Containing nouns and adjectives:

4   noun, adjective, etc. + noun + *aka*-suffix forming bahuvrīhis: *šir-nāmē* (M šyr-n'my 'good' + 'name') 'famous' (cf. *širnām* 'fame'), *δwāts-rītē* 'having twelve faces (*rīt*)', *ān-kutrē* 'originating from the An family (*kutr-*)'.

6.4 Collocation

6.4.1 Hendiadys

Two near synonyms deriving from the same root are combined to make a kind of dvandava compound, occasionally with a conjunction *ət(i)* 'and'. They are called etymological hendiadys and are quite common in Sogdian. Of the two members, the shorter form precedes. When they consist of the same number of syllables, the one beginning with vowel precedes: *wazti frawazti* 'he flies about' < *waz-* 'fly', *frī āfrī* (pryh ʾʾpryh) 'very dear', *ankrantt ət frakrantt* 'he cuts and chops' < *krant* 'cut', etc. Synonymous hendiadys, usually asyndetic, are no less common: *yōk fsāk* 'teaching (and) training', *nom paδka* 'law and judgement', *ramē γēθ* 'cattle flock'. However, Sogdian lacks such hendiadyses as those abundant in Uighur which consist of an Indian loanword and its Uighur synonym: *čakir tilgän* 'wheel' < Skt. *cakra-* (see Gabain 1974: 159).

6.4.2 Phrasal verbs or idioms

Idiomatic combinations or phrasal verbs are not uncommon in Sogdian. Examples are: *θβār θβar-*, lit. 'gift give', = 'give', *rāθ šaw-*, lit. 'road go' = 'travel', *namāč βar-*, lit. 'honour bring' = 'pay homage', *waδu parēp*, lit. 'wife take' = 'get married', *zāwar θβar-*, lit. 'power give' = 'help'. While all these are transparent and are easily analysable, the origin of

kūmārī parēs (kwm'ry prys) 'understand perfectly' seems to have been already opaque to the Sogdians. It is a combination of *kū mārī parēs*, lit. 'arrive at memory (*mār* 'spell')'.

7 SOCIOLINGUISTIC ASPECTS

7.1 Polite phraseology and signals

7.1.1 Honorific expressions

7.1.1.1 Verbs

Sogdian possesses several ways of expressing respectful and humble feelings on the part of a speaker. The honorific expression for 'say' is *framāy*, lit. 'to order' while the opposite is expressed by *patškway* and the unmarked lexeme is *wāβ*. Thus when the Buddha or a king says, *framāy* is employed whereas disciples or servants are *patškway*-ing:

> rti   nūkar əxu ānant kū βayān βaxtam buti   mā0 patīkway . . .
> and now   the A.   to devātideva   Buddha   thus said . . .
>   rti   əxu βayān βaxtam buti   əwən ānant mā0 frāmāy . . .
>   and the devātideva   Buddha the   A.   thus said . . .
> 'then (the disciple) Ananda said to the devātideva Buddha . . .
>   the devātideva Buddha said to Ananda . . .'.

The corresponding nouns *framān* and *patškwān* mean '(His Majesty's) word' and 'humble message' respectively.

The verb *framāy* combined with the past infinitive is also used as an honorific expression meaning 'deign to . . .'. The opposite is the verb *raxn-* 'venture to do':

> framāy   əti-mī   ōnō wiδβāy   əkte
> order.IMPV COMP-me that explanation make.INF
> 'please explain that to me'.

> əzu čānō raxnu   əti-n   yunē əsfrāčk əpsu
> I   how dare.INJ.1s COMP-PART this thing   ask.INJ.1s
> 'How dare I ask (the Buddha) about this matter?'

Notice that in the sentence cited above, politeness is also expressed by the injunctive forms *raxnu* and *əpsu*. On the hypothetical particle *-(a)n* see section 3.2.3.1f.

The verb *xrām* 'walk' seems to have a shade of honour, i.e. 'walk with dignity':

> rti-šī   mā0 patīškway xrām əti   βaya . . .
> and-him thus said   walk COMP lord!
> 'he said to him (the prince): come, O Lord!, . . .'.

7.1.1.2 Nouns and other expressions

Pleonastic use of the vocative form *βaya* 'O Lord!, sir' or its enclitic counterpart *-βay* is one of the most common ways of expressing one's feeling of honour towards the addressee:

> rti-šu   mā0 patīškwayant zārī əti   βaya sayēm
> and-himthus said.3p   sad COMP lord! seem.1p

    pār-ti   *βaγa māx əmī uxušu βrēwar īnč*   *anyu zātē nēst*
    for-COMP lord! us the 6   10,000 woman other son is.not
'they said to him: we feel sad, O Lord!,
    because, O Lord!, for us 60,000 women there is no other son'.

    rti-βaγ   *əzu maδe əkū čāčnīk*   *xūβ ēsu* *rti-βaγ*   *uβyu pōstēt patīwēδu . . .*
    and-lord! I   here to of.Chach king came and-lord! both letters submitted . . .
'O Lord! I came here to the king of Chach. O Lord! I handed both the letters . . .'.

farn (S prn) 'fortune, majesty' is also sometimes used as an honorific word, compare *šmāxfarn* 'Your Majesty', *tawafarn* 'Your (sg.) Majesty'.

As in New Persian 2p forms instead of 2s may be employed for expressing honour:

    Dāwiδ sūkβār farn-sār   *γarf žām*   *namāč . . . anβarz βarta . . .*
    D.   monk majesty-toward many humble homage . . . care   bring.IMPV.2p . . .
        əzu šmāx   *βantē*   *im*
        I   your.PL servant am
'to David the monk. much humble homage to (Your) Majesty . . . (please) pay (2p) heed to . . . I am your (2p) servant'.

7.1.1.3 Expressions showing the speaker's humility

In the passage cited in the last section one finds other means of expressing one's feeling of humility: *žām* 'humble', lit. 'fine' and *βantē* '(your) servant'. *βantē* 'male servant' or *δāy* 'female servant' is a humble expression for 'I' and is often accompanied by *kaštar* 'less' or other synonyms:

    pīšt čan xēp0 kaštar satu βrēwarmīk βantē
    sent from your less   100 10,000th   servant
    'sent from your most trifle one millionth servant'.

Notice that among the elements discussed above *patškway* (> *patškwān*), *raxn-*, and injunctive forms are also counted as showing humility on the part of a speaker or addresser.

7.2 Dialects and variations

7.2.1 Archaic and late features

7.2.1.1 Ancient letters: oldest stage

The oldest and latest strata of the Sogdian language differ considerably from each other. The oldest stage is represented by the so-called 'Ancient Letters'. One finds several features peculiar to the material. These include:

1   Some nouns that show special inflections totally unknown in later Sogdian:
    inst.-abl. *δuγθrya* (δwγδryh) < *δuγta* 'daughter',
    gen.-dat. *māθrī* (m'δryh) < *māt* (m'th) 'mother'.
2   The abundance of the ergative construction which is noticeable because the construction is later displaced by the transitive preterite with the auxiliary *δār*:

> əst Ōrē sarð čānō-ti-m̱-(a)n-č   pačβant əβ̱yart
> is 3   year since-COMP-<u>me</u>-them-from answer <u>obtained</u>
> 'It is three years since I have received an answer from them.'

As regards the phonology, several light stem forms are without vocalic endings: e.g. *əβiž* ('βyz) 'evil', *kərt* (krt) 'made', etc., but it is hard to assess this fact in the history of Sogdian phonology.

7.2.1.2 Christian Sogdian text C5: latest stage

The latest stage of the Sogdian language is represented by Christian Sogdian manuscript C5. At this stage:

1   No article is employed.
2   The conjunction *rti* never occurs and its function of connecting sentences is displaced by *ət*.
3   The conjunction *kat* 'that' plays the function of relative particle as well.
4   The nominal inflection shows strong tendency to use nominative form in all the case functions. Thus in C5 only *x-* and *y-*forms of the extended demonstratives are attested: *xēð, xānt, ēð, yānt.*
       On the other hand:
5   The new oblique case begins to be formed by adding *-ī* to nom. sg. and to similar generalized forms. Compare the inflexions of *ram-* (masc.) 'people' and *wan-* (fem.) 'tree': sg. dir. *ram-i*, obl. *ram-i-(y)ī*; pl. dir. *ram-ta*, obl. *ram-ta-(y)ī*; sg. dir. *wan-a*, obl. *wana-(y)ī*.

Had the inflection been extended to all nouns, the language would have become a fully agglutinative type like Yaghnobi, the so-called 'Modern Sogdian' (see Sims-Williams 1982: 69–70).

7.2.2 Sogdian dialects?

Linguistic differences observed within the written materials are relatively trivial in view of the wide area where Sogdian texts were discovered. Most differences can be explained as chronological or orthographic. For example the durative particle *skun* appears in the forms ('skwn, skwn, C sqn, sk'n, kn (C qn), sk, kw, and k). ('skwn) and (skwn) which are mainly attested in Buddhist and Manichaean texts are older than (sqn, kn, kw), and (k) found in Christian and other late texts and one can assume that gradual reduction of unaccented particles is reflected in the spellings: *əskun* > *skun* > *(s)kən* > *(s)kə* > *(s)k*. A similar reduction of the future particle *kām* is observed.

Henning once argued that the linguistic difference between Manichaean and Christian texts would most probably be due to the sociolinguistic factors (see Henning 1958: 105–108). According to him, in Turfan where the bulk of Manichaean and Christian Sogdian texts were discovered, the Manichaean church was supported by the aristocratic believers whereas ordinary people were the members of Christian church, and the differences between the two varieties are due to the social stratification of the both groups of speakers. Although Henning was reluctant to admit the existence of Sogdian dialects reflecting geographical distribution, one may be reminded that in Christian texts the two forms representing the durative past, i.e. *āz*-imperfect and *ēk(w)māt* imperfect, show

complementary distribution, and that the fact may most reasonably be explained by supposing two dialects which developed different forms for a single category (see Yoshida 1980). However, the difference is trivial and may not deserve to be called dialectal.

The Sogdian language documented by the materials known to us most probably represents the standard variety spoken in Samarqand. A Chinese pilgrim Xuanzang (?-664) who went to India via Sogdiana reported that the manners of Samarqandians were imitated by other Sogdians (see Beal 1884: 33). In fact the Bukharan dialect as cited by Islamic writers is slightly different (see Sims-Williams 1989b: 165–166).

7.2.3 Sogdian and Yaghnobi

The wide range of linguistic difference which once existed in Sogdiana may be inferred by comparing Sogdian with Yaghnobi.

1   The absence of the rhythmic law in Yaghnobi may be explained by its loss in the course of development.
2   However,

    (a) the formation of the imperfect stem by adding the augment *a-* to any present stem: Yaghn. *piraxs-* > *a-piraxs-* 'left' vs. Sogd. *parəxs-* > *pārəxs*, and
    (b) the 3rd pl. ending *-or*:
        Yaghn. *a-wen-or* 'they saw' vs. Sogd. *wēn-ant*, cannot be due to the linguistic change.

It is worth noting that the *-r* ending of 3p is shared by the neighbouring Choresmian and Khotanese languages.

7.3 Language contact and loanwords

The land of Sogdiana was surrounded by other Iranian-speaking areas such as Bactria (to the south), Chorasmia (to the west), and Khotan (to the east). The three languages show a strong affinity to Sogdian and they share several areal features (see Sims-Williams, 1983a and idem 1989a: 169–170). However, no loanwords borrowed from Choresmian and Khotanese has so far been noticed, whereas a few Bactrian elements are known in Sogdian: *saxr* (sxr) 'wheel' and *raxā* (rx'kh) 'cart'.

7.3.1 Borrowings

7.3.1.1 Contact with Bactrian or Kushan influence

Bactrian, the national language of the great Kushan empire, apparently had prestige over Sogdian, because the two forms just cited had cognates in Sogdian, i.e. *čaxr* (cxr) and *raθ-* (rδδ-), and may have been borrowed as culture words (see Sims-Williams 1996a: 50–51). Compare also *paθfar-* (pδβr-) 'honour, rank' borrowed from Bact. πιδοφαρο [piδfar] 'honour' (see Sims-Williams 2004: 541). Similarly *sārtpāw* (s'rtp'w) 'caravan-leader' is suspected to be a loan from Bactrian and this form consisting of Indian *sārt* (cf. Skt. *sārtha-* 'caravan') and Iranian *-pāw* induces one to assume that some loanwords of Indian origin were borrowed into Sogdian via Bactrian. One possible example is *pani* (pny) 'small copper coin' ultimately going back to Skt. *paṇa*, which was borrowed into Bactrian as πανα.

Other Indian elements in Sogdian may also be due to the Kushan influence. Thus Indian loanwords attested in the Ancient Letters belong to the field of transaction, such as *mīδi* (mwδy < Skt. *mūlya*) 'price', *sārt* (< Skt. *sārtha*) 'caravan', *prastak* (< Skt. *prastha* 'pint (measure of capacity)'), etc., and the Sogdians may have learnt them from the Indian merchants travelling through the territory of the Kushan empire. Moreover, several Greek elements may also have come from Bactrian: *δiδəm* (δyδ(y)m < διαδημα 'diadem', *nom* (nwm < voμoς 'law', *δraxm-* (δrxm < δραχμη drachma', *kapiθ* (M qpyδ < καπηλειον 'shop', *miδamβan* (< μεδιμνος 'measure of capacity)', etc. Of course they could equally be a legacy of Hellenism in Iranian languages in general.

7.3.1.2 Loanwords from Western Middle Iranian

Strong cultural influence came from Parthia and Sassanian Iran, and Western Iranian loanwords are numerous. For example all the words for days of a week are of Middle Persian origin: *mīr* 'Sunday', *māx* 'Monday', *wanxān* 'Tuesday', etc. (see Henning 1937: 85–86). A Middle Persian word *rōč* 'day' is usually added to the Sogdian name of a day of a month, e.g. *xurmazt rōč* 'day of Xurmazt, i.e. the first day of a month'. *mayδaβ-* 'minister' is from Parthian and is once accompanied by another Parthian word *wazark* 'great': wz'rktt mγδβt' 'great ministers'.

Vocabulary concerning Christianity is also from Western Iranian: *tarsāk* 'Christian', *sūkβār* 'monk', *masiδar* 'presbyter', etc. (see Benveniste 1964).

However, the alleged contact with Old Sogdian and Old Persian has been refuted as illusory (see Sims-Williams 1989a: 171).

Some Indian elements came via Manichaean Parthian, e.g. *čaxšāpat* (< Parth. cxš'byd < Skt. *śikṣāpada*) 'precept', *bāšīk* 'hymn' (< Parth. b'šh < Skt. *bhāṣā*).

7.3.1.3 Indian loanwords

However, not all the Indian elements came via Parthian. Thus, such words as *makara* 'monkey' (mkr' < Skt. *markaṭa*), *nāk* 'dragon' (n'k < Skt. *nāga*), *ratn-* 'jewel' (rtn- < Skt. *ratna*), *šaman-* 'Buddhist monk' (< *ṣamana* < Skt. *śramaṇa*), etc. which are fully naturalized in Sogdian may represent Indian cultural influence in general. Some phonological features found in those words are proved to be of North Western Prakrit (generally known as Gāndhārī) origin (see Sims-Williams 1983a).

Naturally, Indian elements abound in Buddhist texts. They are of three groups: (a) loanwords naturalized in Sogdian: e.g. *samutr-* (< Skt. *samudra*) 'ocean', (b) Buddhist terminologies well established in Sogdian: e.g. *bodisatβ* 'Bodhisattva' (< Skt. *bodhisattva*), and (c) nonce borrowings employed by certain translators: e.g. *prātimokš* (< Skt. *prātimokṣa* 'moral code').

7.3.1.4 Borrowing from other languages

As the traders of Silk Road, a number of Sogdians were resident in China and one might expect many borrowings from Chinese. However, they are not at all numerous, and a very few which actually found their ways into Sogdian are attested also in other Central Asian languages, such as Uighur and Tocharian (see Yoshida 1994: 379). For example a word for 'pint' *šang* which comes from Chinese 升 (Middle Chinese **siəng*), also found in Uighur *šing*, Tocharian *ṣank*, and Khotanese *ṣamga*, *śimga*. The situation suggests that they were culture words widely in use among the peoples of Central Asia, and that

they were not necessarily borrowed directly from Chinese into Sogdian. Similarly, the languages of Turkish peoples, with whom Sogdian had close contact, lent very few words to Sogdian. One of the possible examples, other than personal names and titles, in late texts is *arxiš* 'caravan' (< Uighur *arqīš*) which seems to have displaced a Indian counterpart *sārt* by that time.

Tocharian elements are also very few. One example *sintāp*, which occurs in a medical text, seems to be borrowed from Tocharian B *sintāp* 'a kind of rock-salt', ultimately derived from Indian, compare Skt. *saindhava* 'rock salt'.

Recently, *čabiš* (cpyš), which alternates with an inherited word *sarθang* (srδng) 'general (of army)', has been argued to have originated from the language of the Hephthalites. The Hephthalites ruled Sogdian for some hundred years beginning in the latter half of the fifth century and their language could have influenced Sogdian, though no other Hephthalite words have been known.

7.3.2 Sogdian elements in other neighbouring languages

In stark contrast with the borrowings, Sogdian lent many words to Uighur, an Old Turkish dialect, e.g. *ažun* (< *āžun*) 'existence', *känt* (< *kanθ*) 'town', *nizwani* (< *nizβānē*) 'sin', *tamu* (< *tam-*) 'hell', etc. Sogdian's prestige over Uighur was such that the latter adopted Sogdian script for writing Uighur. A so far unnoticed Sogdian element in other Turkic languages than Uighur is *axšam* 'evening' which goes back to *əxšām* ('xš'mh) 'evening meal'.

The influential Sogdian culture and language in the early Islamic period may account for relatively numerous Sogdian loanwords in New Persian (see Henning 1939): e.g. *čayz* 'frog' < *čayz* (cyz), *čuɣd* 'owl' < *čɣut* (cɣwt), etc.

The reason why virtually no Sogdian loanword is found in Chinese may be due not only to the latter's prestige but also to its natural reluctance to borrowing foreign words in general. Only two forms which found their way into dictionaries are *chi bo* 叱撥 (Middle Chinese *tšʻiet puat*) denoting a special kind of horse which is derived from *čarθpāδ* (cyrδp'δ) 'quadruped' and *mi* 蜜 (Middle Chinese *miet*) from *mīr* 'Sunday (ultimately from Middle Persian *mihr*)'.

7.4 Language death

The latest known Sogdian text is dated to 1025 CE (see Livšic 1996: 271 and Yoshida 2004: 21). Possibly within a few centuries after that the Sogdian language was no longer spoken or written and became a dead language.

7.4.1 Death of Sogdian: Turkicization and Persianization

7.4.1.1 Turkicization

The changing relationship between Sogdian and Uighur is reflected in several late texts where many Turkish elements (words, personal names, expressions calqued on Uighur) are found. A text discovered in Dunhuang and dated to the tenth century is bilingual in Sogdian and Uighur; in it the scribe alternates between writing in Sogdian and in Uighur, and it is difficult to find reasons for the use of one or the other (cf. Sims-Williams and Hamilton 1990: 24–25):

Uighur

temči-dä   bir   qïzïl   qars   alyu   bar
Temči-by   one   red   wool   to.be.taken is
'by Temči one   red   qars   is to be taken'.

Sogdian

alp irkinī   nəβant əšē   karmīr rayzi āsī   xu
Alp Irkin.OBL by   three red   wool to.be.taken is
'by Alp Irkin three red rayzs are to be taken'.

Sogdian appears to have been in the process of being absorbed into Turkish first in the eastern part of the Sogdian speaking area. Maḥmūd al-Kāšyarī of the eleventh century reports that in his time people from Soyd (area between Bukhara and Samarqand) were resident in Balāsāyūn (a city in modern Semirechie, Kirghiz), and that their dress and manner was that of Turks (see Dankoff and Kelly 1982: 352 and Yoshida 2009). The description of those Sogdians may point to their profound Turkicization in the eleventh century.

7.4.1.2 Persianization

The examples of the languages that al-Muqaddasī in the tenth century reports as those of Samarqand (*be-goftag-om* 'I have said') and Bukhara (*dānest-ī* 'you.2s knew') are nothing but dialects of New Persian (see al-Muqaddasī 1994: 335), and show that Sogdian was no longer in use there by that time. This conclusion is also reflected in a note of Iṣṭakhrī of the tenth century that in his time the inscription on the Kish gate of Samarqand was incomprehensible to the local people. It is likely that the inscription, which Iṣṭakhrī took for Himyaritic, was in Sogdian (see Barthold 1958: 87). On the other hand, al-Muqaddasī also mentions a vernacular spoken in Suyd (the area lying between Bukhara and Samarqand) which sounded similar to a dialect found in the suburbs of Bukhara.

This suggests that during his lifetime Sogdian was still spoken in small villages, while New Persian was spoken in the urban areas. Today, the sole survivor of the Sogdian dialects, once spoken all over Sogdiana, is Yaghnobi, which continues to be spoken in the most remote valley of the Yaghnob (North Tajikistan).

8 SAMPLE TEXT

The following sample texts reproduce the main part of the Manichaean Sogdian 'Story of the Pearl-Borer', pp. 466–469 in Henning, W. B. (1945). The text is in two scripts, one in Manichaean script and the other in Sogdian. (Round brackets) indicate damaged letters; [square brackets] indicate suggested restorations of missing letters (see Henning, *ibidem*). The text is arranged here by clause divisions. For each clause there are five lines:

1   transliteration of Manichaean script;
2   transliteration of Sogdian script;
3   transcription;
4   literal interlinear translation;
5   idiomatic translation, mostly following Henning.

1

'rtyxw	mr'z	mrtyy	kw	xtw	s'r	w'nw	ptyškwy
[rty ZK]	mr'z		kw	'xtw	s'r	(w'n')[kw	ptyš]kw'y
(ə)rti xu	marāz	martī	kū	əxtu	sār	wānō	patīškway
and the	hired.labour	man	to	judge	towards	thus	said.IMPF

'The hired labourer addressed the judge thus:

2

ktβγ	ywnyy	fšy'ws	c'nw	t'm'		w'crn	kwsyy	wyn
ktβγ	ywn'k	'βšy-'ws	c'n'kw	t'm'k	ZKwy	w'crn'y	kwsy	wyn
kat-βay	yunē	(ə)fšəyāus	čānō	tāmā	əwī.LOC	wāčarn(ī)	kōsī.OBL	wēn.IMPF
that-lord!	this	gentleman	when	me	the	bazaar	side	saw

"My lord, when this gentleman saw me at the side of the bazaar,

3

'tmyy	w'nw	ps'	kδwtyh	cw	'rq	γrβyy
rtmy	w'n'kw	ps'	'YKZY	cw	'rkh	γrβ(')[y]
ərt/ət-mī	wānō	psa	kaδ-uti	ču	ark	γərβe
and-me	thus	asked	that	what	work	you.know.2s

he asked me: "Hey, what work can you do?"

4

'rtšy	'zw	w'nw	ptyškwyy	kt	βγ	wyspw	'rk
rtšy	'zw	w'n'kw	ptyškw'y	[k](t)	βγ	wysph	'rkh
(ə)rt-šī	əzu	wānō	patīškway	kat	βay	wyspula.ACC	ark
and-him	I	thus	said	that	lord!	all	work

'I replied: "Sir, whatever work

5

cw'tmy	tγw	frm'yy	'rty	s't	wyspw	'rk	γrβ'm
cw-[ZY](m)y	tγw	pr'm'yy	rty	s't	wys(p)[h]		γrβ'm
ču-ət-mī	tayu	framāye	ərti	sāt	wispula	ark	γərβām
that-COMP-me	you	order	and	all	every	work	I.know

you may order me (to do), I can do it all" '.

6

'tymy	c'nw	kw	x'n'	s'r	šykr
rtmy	c'n'kw	kw	x'n'kh	s'r	šykr
ərt-mī	čānō	kū	xānā	sār	šīkar
and-me	when	to	house	towards	lead.IMPF

'When he had taken me to his house,

7

'rtmy	wyn'	fr'm'y	jtyy
rty my	wyn'kh	pr'm'y	z-yt'y
ərti-mī	wīnā	frāmāy	žəte
and-me	lute	ordered	play.INF

he ordered me to play on the lute'.

8

’rty	βy’ryy	prm	prw	xypδ’wndy<u>h</u>		frm’n	wyn’	jy<u>t</u>(w) δ’rm
rty	βy’r’k	prm	prw	xypδ’w’nty		prm’nh	wyn’k	z-ytw δ’r’m
ərti	*βyārē*	*parm*	*paru*	*xēpθāwantī*.OBL		*framān*	*wīnā*	*žətu-δāram*
and	evening	until	on	owner’s		order	lute	play-PRET.1s

‘Until nightfall I played on the lute at the owner’s bidding’.

9

’rt̠šw	x<u>t</u>w	w’nw	pδk’	nym’y
rtšw	’xtw	w’n[kw]	pδkh	nym’y
ərt-šu	*(ə)xtu*	*wānō*	*paδka*	*nīmāy*
and-the	judge	thus	verdict	judge.IMPF

‘The judge pronounced this verdict:’

10

k<u>t</u>	<u>t</u>γw	wny	mr<u>t</u>y	mr’z	p<u>t</u>xry<u>t</u>δ’ryy
’ỾKZY	tγw	ZKn	[m](r)ty	mr’z	ptxr’ytδ’r’y
kat	*taγu*	*wənī*	*martī*	*marāz*	*patxrīt-δāre*
that	you	the	man	hired.labour	hire-PRET.2s

“You contracted that man’s (skilled) labour”
(which includes musical performance).

11

’rt̠šw	cqn’c	pyδ’r	mrγ’r<u>t</u>	nyy	frm’yy	swβ<u>t</u>y
rtβ	[ckn](c)w	pyδ’r	mrγ’rt	L’	pr(’m)[’yy	swβty
ərt-šul-β	*čakanāč*	*piδār*	*marγārt*	*nē*	*framāye/frāmāye*	*suβte*
and-him.PART	from.what	sake	pearl	not	order.PRES.2s	bore.INF
					ordered.IMPF.2s	

“So why did you not order him to bore the pearls,

12

p’rtyšy	xww	wyn’	jnyy	frm’<u>t</u>δ’ryy
[p’rZYšy	ZK	wyn’kh	z-n’y	prm’t δ] ’r’y
pār-ti-šī	*xu*	*wīnā*	*žani*	*framāt-δāre*
but-COMP-him	the	lute	play.INF	order-PRET.2s

“but bid him play on the lute instead?”

13

’rty	wny	mr<u>t</u>y	xww	prxyy	’spty<u>h</u>	δβ<u>t</u>yy	γw<u>t</u>-k’m
rty	(ZKn)	[mrty	ZK]	prxy	’spt’k	δβ’rt’k	γwt-k’m
ərti	*wənī*	*martī*	*xu*	*parxē*	*əsptē*	*θβartē*	*γawt-kām*
and	the	man	the	wage	completely	given	needs-FUT

“That man’s wages will have to be paid in full.”

ACKNOWLEDGEMENT

BIBLIOGRAPHY AND REFERENCES

Barthold, W. (1958) *Turkestan down to the Mongol Invasion*, 2nd ed., London: Luzac & Co.

Beal, S. (1884) *Si-Yu-Ki. Buddhist Records of the Western World*, Vol. 1, London: Trübner.

Benveniste, E. (1954) 'Expression de «pouvoire» en iranien', *BSL* 50: 56–67.

——(1963 [1964]) 'La vocabulaire chrètien dans les languages d'Asie Centrale', in *Atti del convegno internaz sul tema: L'Oriente cristiano nella storia della civiltà*, Rome: Academia Nazionale dei Lincei, pp. 85–91.

——(1966) 'Le verbe *stā-* comme auxiliaire en iranien', *AO* 30: 45–49.

Dankoff, R. and Kelly, J. (1982) (eds) *Maḥmūd al-Kāšɣarī: Compendium of the Turkic Dialect*, Part 1, Cambridge: Harvard University Printing House.

Gabain, A. von (1974) *Alttürkische Grammatik*, 3rd ed., Wiesbaden: Otto Harrassowitz.

Gershevitch, I. (1945) 'Sogdian compounds', *TPhS*: 137–149.

——(1954) *A Grammar of Manichean Sogdian*, Oxford: B. Blackwell.

Gharib, B. (1965) *Analysis of the Verbal System in the Sogdian Language*, PhD dissertation, University of Pennsylvania.

——(1995) *Sogdian Dictionary*, Tehran: Farhangān.

Grenet, F. and Sims-Williams, N. (1987) 'The historical context of the Sogdian Ancient Letters', in *Transition Period in Iranian History, Actes du symposium de Fribourg-en-Brisgau (22–24 Mai 1985)*, Leuven: Peeters, pp. 101–122.

Henning, W. B. (1936 [1937]) *Ein manichäisches Bet- und Beichtbuch, APAW* no. 10.

——(1939) 'Sogdian loan-words in New Persian', *BSOS* 10: 93–106.

——(1945) 'Sogdian tales', *BSOAS* 11: 465–87.

Heston, W. L. (1976) *Selected Problems in Fifth to Tenth Century Iranian Syntax*, PhD dissertation, University of Pennsylvania.

Livšic, V. A. (1996) 'Glave 7: Sogdijskie teksty, dokumenty i epigrafika', in Ploskij, V. M. *et al.* (eds) *Istočnikovedenie Kyrgyzstana (c drevnosti do XIX v.)*, Biškek: Ilim, pp. 215–273.

Lubotsky, A. (2002) 'Scythian elements in Old Iranian', in Sims-Williams, N. (ed.) *Indo-Iranian Languages and Peoples*, Oxford: Oxford University Press, pp. 189–202.

al-Muqqadasī (1994) *Aḥsan al-taqāsīm fī ma 'rifat al-aqālīm* (The best divisions for knowledge of the regions), Collins, B. (trans.), Reading: Centre for Muslim Contribution to Civilization Garnet Publishing.

Paul, L. (1997) 'Präteritum und Perfekt im Soghdischen', *IF* 102: 199–205.

Provasi, E. (1997) 'Note sulle construzioni relative in Sogdiano', in Ambrosini, R. *et al.* (eds) *Scribthair a ainm n-ogaim. Scritti in memoria di Enrico Campanile*, Ospedelatto, Pisa: Pacini, pp. 815–839.

Sims-Williams, N. (1973) 'Notes on Bactrian syntax', *IF* 78: 95–99.

——(1979) 'On the plural and dual in Sogdian', *BSOAS* 42: 337–346.

——(1981) 'The Sogdian sound-system and the origins of the Uyghur script', *JA* 269: 347–360.

——(1982) 'The double system of nominal inflexion in Sogdian', *TPhS* 1982: 67–76.

——(1983a) 'Chotano-Sogdica', *BSOAS* 46: 40–51.

——(1983b) 'Indian elements in Parthian and Sogdian', in Röhrborn, K. and Veenker, W. (eds) *Sparachen des Buddhismus in Zentralasien*, Wiesbaden: In Kommission bei O. Harrassowitz, pp. 132–141.

——(1984) 'The Sogdian "rhythmic law" ', in Skalmowski, W. and. van Tongerloo, A (eds) *Middle Iranian Studies. Proceedings of the International Symposium Organized by the Katholitheke Universiteit Leuven from the 17th to the 20th of May 1982*, Leuven: Peeters, pp. 203–215.

——(1985a) *The Christian Sogdian Manuscript C2*. (Schriften zur Geschichte und Kultur des alten Orients. Berliner Turfantexte XII), Berlin: Akademie-Verlag.

——(1985b) 'A note on Bactrian phonology', *BSOAS* 48: 111–116.

——(1986a) 'Sogdian "δprm and its cognates', in Schmitt, R. and Skjaervø, P. O. (eds) *Studia Grammatica Iranica. Festschrift für H. Humbach*, München: R. Kitzinger, pp. 407–424.

——(1986b [1987]) "Sogdian *kw* and Slavonic *ku*', *Peredneaziatskij Sbornik* 4: 116–21.

——(1988) 'Syro-Sogdica III: Syriac elements in Sogdian', in Sundermann, W. *et al.* (eds) *A Green Leaf. Papers in Honour of Professor Jes P. Asmussen* (Acta Iranica 28. Hommage et Opera Minora, XII), Leiden: Brill, pp. 145–156.

——(1989a) 'Sogdian', in *CLI*, pp. 173–192.

——(1989b) 'Eastern Middle Iranian', in *CLI*, pp. 165–172.

——(1990) 'Chotano-Sogdica II: Aspects of the development of nominal morphology in Khotanese and Sogdian', in Gnoli, G. and Panaino, A. (eds) *Proceedings of the First European Conference of Iranian Studies*, Part 1, Roma: Instituo Italiano per il Medio ed Estremo Oriente, pp. 275–296.

——(1994) 'The triple system of deixis in Sogdian', *TPhS* 92(1): 41–53.

——(1996a) 'The Sogdian merchants in China and India', in Cadonna, A. and Lanciotti, L. (eds), *Cina e Iran*, Firenze: Leo S. Olschki, pp. 45–67.

——(1996b) 'On the historic present and injunctive in Sogdian and Choresmian', *MSS* 56: 173–189.

——(1996c) 'The Sogdian manuscripts in Brāhmī script as evidence for Sogdian phonology', in Emmerick, R. E., Sundermann, W., Warnke, I. and Zieme, P. (eds) *Turfan, Khotan und Dunhuang. Vorträge der Tagung 'Annemarie v. Gabain und die Turfanforschung', veranstaltet von der Berlin-Brandenburgischen Akademie der Wissenschaften in Berlin (9.–12. 12. 1994)*, Berlin: Akademie Verlag, pp. 307–315.

——(2000) *Bactrian Documents I*, London: Nour Foundation in association with Azimuth Editions and Oxford University Press.

——(2004) 'The Parthian abstract suffix *-yft*', in Penny, J. H. W. (ed.) *Indo-European Perspectives. Studies in Honour of Anna Morpurgo Davies*, Oxford: Oxford University Press, pp. 539–547.

——(2007) 'The Sogdian potentials', in Macuch, M. et al. (eds) *Iranian Languages and Texts from Iran and Turan. Ronald E. Emmerick Memorial Volume*, Wiesbaden: Harrassowitz, pp. 377–386.

Sims-Williams, N. and Grenet, F. (2006) 'The Sogdian inscription of Kultobe', *Shygys* (Almaty) 2006.1 [2007]: 95–111.

Sims-Williams, N. and Hamilton, J. (1990) *Documents turco-sogdiens du IXe–Xe siècle de Touen-houang*, London: Published on behalf of Corpus Inscriptionum Iranicarum by SOAS.

Telegdi, S. (1938) 'Notes sur la grammaire du sogdien chrétien', *JA* 230: 205–233.

Weber, D. (1970) *Die Stellung der sog. Inchoativa im Mitteliranischen*, PhD dissertation, Göttingen.

—— (1971) 'Also sprach Mandrī (Bemerkungen zur direkten Rede im Sogdischen)', *IF* 76: 77–83.

—— (2000) 'Zur grammatisch-semantischen Bestimmung einiger Lemmata des sogdischen Lexikons', in Emmerick, R. E. *et al.* (eds), *Studia Manichaica IV. Internationaler Kongreß zum Manichäismus, Berlin, 14.–18. Juli 1997*, Berlin: Akademie Verlag, pp. 639–647.

Yakubovitch, I. (2005) 'The syntactic evolution of Aramaic ZY in Sogdian', *StIr* 34(2): 199–230.

Yoshida, Y. (1979) 'On the Sogdian infinitives', *Journal of Asian and African Studies* 18: 181–195.

—— (1980) 'On the dialectology of Christian Sogdian [in Japanese]', *Oriento* 23(1): 83–93.

—— (1994) 'Chinese in Sogdian script [in Japanese]', *The Toho Gakuho* 66: 271–380.

—— (2000) 'First fruits of the Ryukoku-Berlin joint project on the Turfan Iranian manuscripts', *Acta Asiatica* 78: 71–85.

—— (2004) 'Dating of linguistic changes encountered in the texts unearthed from the Silk Road [in Japanese]', *Ex Oriente* 11: 3–34.

—— (2005) 'The Sogdian version of the new Xi'an inscription', in Trombert, E. and de la Vaissière, E. (eds) *Les Sogdiens en Chine*, Paris: École française d'Extrême-Orient, pp. 57–72.

—— (2006) 'Honorific and polite expressions in Sogdian', in Shirai, S. and Shogaito, M. (eds) *Philological Studies on Old Central Asian Manuscripts*, Kyoto: Department of Linguistics, Kyoto University, pp. 81–94.

—— (2009) 'Turco-Sogdian features', in Sundermann, W., Hintze, A. and de Blois, F. (eds), *Exegisti monumenta. Festschrift in honour of N. Sims-Williams*, Wiesbaden, Harrassowitz, pp. 571–585.

Yoshida, Y. *et al.* (1989) 'A Sogdian sales contract of a female slave from the period of the Gaochang Kingdom under the Rule of Qu Clan [in Japanese]', *Studies on the Inner Asian Languages* IV: 1–50.