

Introduction

Sogdian language

Sogdian was a Middle Northeast Perso-Aryan language once spoken in the land of Sogdiana, first mentioned in the Avesta, *suγda-*. It was also a *lingua franca* spoken outside, especially, in Chinese Turkeṣtān (Xinjiang Region).

Sogdian scripts

The Archaic Sogdian script was an adaptation of the Achaemenian chancery script which used distinct and separated graphemes. Two cursive scripts evolved from the Archaic script:

- i. One, the “half-cursive” script (also referred to as the formal script, or Sūtra-script)¹, in which the characters are clearly and distinctly written, and some characters are written joined together.
- ii. The other, the “cursive” script², refers to a variety of script styles in which the characters are not clearly written and it is difficult to distinguish one from the other.

The foreign communities also used their own scripts (B: Buddhist Brāhmī; M: Manichaean modified Syriac; C: Christian modified Syriac). Note that the Sogdian scribes employed logograms limited to most common words.

Alphabet

In the Archaic script, a letter of a single word is unconnected from a letter preceding or following it in the same word. In the half-cursive and cursive scripts letters of a single word connect and change shape according to their position within it.

The writing systems: one, horizontally, that is, the direction is from right to left with lines that advance from top to bottom; the other, vertically, that is, the direction was from top to bottom with lines that advance from left to right.

The symbols used in Sogdian script, including both those used in the regular Sogdian vocabulary and those used in logograms, can be found in Table 1. We use both transliteration and transcription. Transliteration consists of converting the Sogdian letters into “Roman” graphemes without any regard to pronunciation, and the number of graphemes in the transliterated string is the same as the number of letters in the Sogdian

¹ . Cf. Pers. *nēmvaštaḡ-dibīrīh*

² . Cf. Pers. *vaštaḡ-dibīrīh*

string. Letters of a logogram are converted to uppercase letters. Transcription consists of writing the phonemes of a Sogdian word in “Roman” letters to show the way it is pronounced. Note that it demands high levels of phonological and morphological knowledge.

Table 1

	Separated	Connected		
		Final	medial	Initial
'	𐰇	𐰇𐰆	𐰇	𐰇
β; B	𐰇𐰆	𐰆	𐰇𐰆	𐰇𐰆𐰆
f			𐰇𐰆	𐰇𐰆
γ; G	𐰇	𐰇	𐰇	𐰇
(d); D	𐰇			𐰇
h	𐰇𐰆	𐰆		
w	𐰇	𐰇	𐰇	𐰇
z	𐰇	𐰇𐰆	𐰇	𐰇
z (ž)			𐰇	𐰇
x; H	𐰇𐰆	𐰇𐰆	𐰇	𐰇
T				𐰇
y	𐰇	𐰇	𐰇𐰆	𐰇𐰆
k	𐰇	𐰇𐰆	𐰇	𐰇
δ; L	𐰇	𐰇𐰆	𐰇	𐰇
m	𐰇	𐰇	𐰇	𐰇𐰆
n	𐰇𐰆	𐰇𐰆	𐰇	𐰇
s	𐰇	𐰇	𐰇	𐰇
'	𐰇𐰆𐰆	𐰇	𐰇	𐰇𐰆
p	𐰇	𐰇	𐰇	𐰇
č; Š	𐰇𐰆	𐰇𐰆	𐰇	𐰇

(q)				𐬑
r	𐬕	𐬖	𐬗	𐬘
š	𐬙	𐬚	𐬛	𐬜
t	𐬝	𐬞	𐬟	𐬠
l		𐬡	𐬢	𐬣

Descriptive phonology

Vowels

Vocalic phonemes are presented thus:

		frontal	central	posterior
high	short	i		u
	long	ī		ū
mid	short	e	[ə]	o
	long	ē		ō
low	short		a	
	long			ā

/a/: <'>, <∅>

/a/ in word-initial position: <'nδm'k> /ándamē/ 'limb, member'; <'nkwšt> /ángušt/ 'finger, digit'; <'sp-y> /asp-í/ 'horse'; <'zw> /azú/ 'I'.

/a/ in word-medial position: <βγ-> /βaγ-/ 'lord, sir, milord (the polite address); god'; <γr-> /γar-/ 'mountain'; <nw''z> /navāz/ 'sailor'.

/ā/: <'>, <'>, <h>, <'kh>

/ā/ in word-initial position: <'p> /āp/ 'water; river'; <'kn> /ākan/ 'to fill'; <'γ'z> /āγāz/ 'beginning'.

/ā/ in word-medial position: <'sm'n> /asmān/ 'sky, heaven'; <'p't> /āpāt/ exclam. 'bravo!'; <δ'y> /δāy/ '(maid-)servant'; <m's'k> /māsē/ 'old man'.

/ā/ in word-final position: <'krtyh> /əkertyā/ 'action, act, deed, work, conduct, kárman'; <β'z'kh> /βāzā/ 'arm'; <xwýčkp'znyh> /x'īck-pāzanyā/ 'open-mindedness'.