LESSON 1

ORTHOGRAHY

THE MANICHEAN SOGDIAN SCRIPT.

The Manichean alphabet is commonly regarded as a variant of the Syriac Estrangelo script, but several letters have shapes closer to the Syriac Nestorian script. Its invention is sometimes attributed to Mani himself, but the alphabet is probably older than that.

The order of the letters in the table below is that of the Aramaic-Syriac alphabets.

<table>
<thead>
<tr>
<th>The Manichean Alphabet</th>
</tr>
</thead>
<tbody>
<tr>
<td>a</td>
</tr>
<tr>
<td>h (ḥ)</td>
</tr>
<tr>
<td>s</td>
</tr>
<tr>
<td>l</td>
</tr>
<tr>
<td>r</td>
</tr>
<tr>
<td>ţ</td>
</tr>
<tr>
<td>j</td>
</tr>
</tbody>
</table>

Notes on the table.

The Syriac letter <l> is used for <δ> and <ς> for Sogdian <ς>.

The letter <δ> is used to write both δ and θ.

The letter <j> is not found in the Syriac version of the alphabet, but is peculiar to Sogdian. In the Middle Persian and Parthian versions of the script a <ژ> with two dots above <ژ> is used instead.

On <اء> (‘ayn) see below.

The letter forms are quite constant in the manuscripts, with the exception of <d, r, ء>, and <k, x>, which vary according to manuscript.

Otherwise, when a letter has two forms in the table, the one to the left is used in final position.

The letters <β>, <γ>, <ς>, and <ς> are modified forms of <b>, <g>, <p>, and <k>.

The letter <γ> (Syriac ḫē) is used only in final position and has no phonetic value, while <γ> <ή> (Syriac ḫē) is found very rarely in loanwords from Parthian (e.g., <κρμšwhn> karmōhan “absolution” Lesson 10). The letter is frequently lengthened to fill the space at the end of a line.

Several letters adjust their forms when there is too little space at the end of a line for their normal forms, e.g., <ς> for <ς>, <ς>, and <ς> for <ς>: <ς> has the special form <ς σωστός>, sometimes at the beginning of words; <ς> has the squeezed form <ς> for <ς>.

Letters with a left extension (<β, γ > etc.) can extend this as much as needed to fill space.

Note also that the letters <ς> and <ς> are usually written inside <ς>.

In double <δδ> the letters are close to one another: <δδ>.
TRANSLITERATION AND TRANSCRIPTION.

When we simply substitute English letters for the Sogdian ones we say we “transliterate” the words, but when we write out the word in English letters the way it was pronounced we say we “transcribe” the words.

To denote that we are simply transliterating we may enclose the transliteration in pointed brackets < >, while transcriptions are indicated by italics.

Example: ʰəmā translit. <m’t >, transcr. mār “mother”; ʰənax <x’n’h > xānā “house”; ʰδyn’r > δēnār “dinar”; ʰyδ<δ > méθ “day.”

VOWELS.

Vowels are not written consistently in the Sogdian scripts, and it is therefore not always certain what they were, although most of the time we can make educated guesses on the basis of orthography and linguistic comparison with other Iranian languages.

As the Sogdian alphabets are of Aramaic-Syriac descent they do not regularly express short vowels in writing. In the Manichean script long vowels are always written, using <ā> for ā, <ē> for ē and ī, and <w> for ō and ū.

The correct vowels have to be learned for each word.

The letter <ā> is used initially (at the beginning of a word) to express a, ə, or long ā, but double <’ā > is commonly written for ā.

At the beginning of a word long ē and ī are written <y-ā> or <y>, while short i and u are written <y- > and <w- >.

ACCENT AND THE RHYTHMIC LAW—LIGHT AND HEAVY STEMS.

Sogdian words consist of a “stem” and an “ending.”

Usually, endings are case endings of nouns, adjectives, pronouns, and adverbs or personal endings of verbs.

The form of a noun, verb, etc., that is left when the “ending” is removed is the “stem.”

A stem may contain one or more suffixes. For instance, əkt- “did” is the past stem of the present stem kun- “does.” With the suffix -yāk it becomes a noun əkt-yāk “act, action,” which is also a “stem.”

The accent in Sogdian lay on the first “long vowel” of the word if it had one. (The nature of “long vowels” will be defined in the next section on vowels.)

If the first long vowel was in the stem, the word was accented on the stem.

If the stem contained no long vowel, the word would be accented on the ending, whether its vowel was short or long.

In this way, all Sogdian words can be characterized as belonging to one of two types. Stems with the accent on the stem are called “heavy stems,” and words with the accent on the ending are called “light stems.”

This system of light and heavy stems is commonly referred to as obeying the “rhythmic law” and affects all Sogdian declensions, conjugations, and word formations. In “heavy stem” words, final short vowels were lost, final long vowels often reduced, and final consonants occasionally lost.

Note: In order to retain important grammatical distinctions short-vowel endings were sometimes restored by analogy with light stems, however.

Example:

Light: ʰβγ- <βγ > “god” nom. sing. ʰβγ-ʰ <βγ-ʰ>

Heavy: ʰβγ-ʾ <βʾγ > “piece of land, garden” nom. sing. ʰβγ-ʾ ʰ <βʾγ>

The “rhythmic law” also affected many suffixes, which took different forms according as the “stem” to which they were attached was heavy or light.

Example:

Light: ʰkt- <kt > “done” ʰkt-yāk “act, action,”

Heavy: ʰγɾbāk “wise” ʰγɾbāk-yā “wisdom”
LESSON 1

VOWEL PHONEMES.

The rhythmic law shows that Sogdian possessed at least the short vowel phonemes /a, i, u/ and the long /ā, ẹ, ī, ū, ő/; a system known from other Iranian languages. These vowel phonemes were probably phonetically realized more or less as “cardinal vowels” when stressed, long and short: [a - ə], [ɛː], [i - iː], [oː], [u - uː].

1. Short vowels.

In this manual short a or ā will be used in transcription of heavy stems, but ə instead of a in the transcription of light stems, in order to enable the student to see at a glance the nature of the stem.

There are basically no Sogdian words ending in a consonant with only short vowels. Any stem of this structure either requires an ending, e.g., /βαγ-/, /βαμ-/, or must be enclitic, e.g., /κατ-βαγ/.

The exact distribution of final short /e, o/ cannot always be verified, as it is not clear which consonant clusters existed. In initial consonant clusters, for instance, (two or more consonants at the beginning of a word) we do not know if vowels were inserted or not. The fact, however, that the orthography in many instances vacillates between nothing and <’< or <’y> indicates that short vowels were sometimes not pronounced.

In these cases, comparison with other languages leads one to posit one or two “central” vowels [ə] (so-called schwa), a vowel like the e in English perhaps, and [i], a vowel sounding like the first e in English between. Thus, <ένας <πτυγ> may have been pronounced (πτγο) in slow and accurate speech, but in normal speech either πτγο <πτγο> or—after a vowel —πτγο, and—after consonant—even πτγο.

This [ə] may also have been influenced by its phonetic context, e.g., before palatal consonant we seem to have [ə] ~ [e] in <έν> βέζ ~ <έν> βέζ or βέζ.

In this manual ə will be used to indicate either of the unstressed vowels [ə] and [i]. The ə is always indicated in the transcriptions, although the principles underlying its inclusion are admittedly impressionistic.

Short /u/ may have been realized as [u], [us], [wa], or [wu] depending on the context. This analysis is based mainly upon the fact that words with original initial Cu- can take a prothetic ə-, e.g., <κυτ> κυτί = [κωτι, κωτι, κωτι]. Other examples are difficult to find.

Similarly, short /i/ may have been realized as [i], [ia], [ya], or [yi] depending on the context.

To simplify the transcription, in this manual u and i will be used, occasionally wə and yə. When ur, ır, un, and in occur in heavy stems they will be marked as stressed: ār, etc.

The exact distribution of final short -i and -e is unclear. Here, certain etymological principles have been followed.

2. Long vowels.

The long vowels /ā, ẹ, ī, ū, ő/ may have been long only in stressed position and short in unstressed position. The variant spellings of the verbal endings may reflect this.

The short /e/ and /o/ were probably not separate phonemes opposed to /ë/ and /ő/. Short [e] seems to be supported by alternances such as in the ending <’ẹn> ~ <-ny>, i.e., *-enē ~ -anē.

There are no similar pairs for [o].

In this manual e is used (e.g., βένδ- “to bind”), but u instead of o (e.g., βουχ- “light,” not βουχ-).

Whether there was an opposition between final stressed /-ē/ and /-ė/, is also very uncertain. In this manual the traditional transcription with final short -ē in some forms of light-stem nouns (adjectives, pronouns) and verbs as opposed to -ē and -ė < *-aka is maintained for pedagogical reasons.

3. Nasalized and rhotacized vowels.

Sogdian apparently had short and long rhotacized (retroflex) and nasalized vowels, phonemically (probably) vowel + [h] or [h].

Not all vowels + [h] produce heavy stems, however. For instance, mury “bird” is a light stem, but mury
“meadow” is a heavy stem. The explanation for the difference is that, historically, the light stems contain an Old Iranian “vocalic r,” which functioned as vowel. Thus, marga “bird,” but marga “meadow.” In Sogdian, the “vocalic r” developed a short vowel before it, which remained short and did not cause a stem to become heavy, while the old sequence vowel + r probably became a long rhotacized /ɾ/. Differently, almost all stems with n before consonant are heavy.

4. Diphthongs.
The existence of short-vowel diphthongs is uncertain, as we have little means of determining whether the old diphthongs ai and au remained before consonants or had become ē and ē. It is possible that they were still diphthongs at an early stage of Sogdian, as suggested by the Sogdian orthography, but were simply long vowels in the stage represented by the Manichean and Christian texts.

In this manual only long vowels ē and ē will be used before consonants, thus *βaw- + -am > βawam, but *βaw- + -i > βoi.

The situation in final position is even less certain.

5. “Long diphthongs.”
The sequences V + i, u, r, n, m (āī, āū, ōr, etc.) are perhaps more conveniently analyzed as combinations of V + y, w, r, n, m, thus /āyC/ = [āywC], /āW/ = [āwʔC], /ānC/ = [ānʔC], etc.

With considerable reservations one may posit the following possible system of (attested) vocalic phonemes and allophones for Sogdian:

<table>
<thead>
<tr>
<th>stressed</th>
<th></th>
<th></th>
<th></th>
<th>unstressed</th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>/ā</td>
<td>[ā]</td>
<td>+ /r/</td>
<td>[a]</td>
<td>+ /r/</td>
<td>[a]</td>
<td>+ /n/</td>
<td>[aŋ]</td>
</tr>
<tr>
<td>/a</td>
<td>[a]</td>
<td>[aɐ]</td>
<td>[aŋ]</td>
<td>/e</td>
<td>[e]</td>
<td>[e̥]</td>
<td>[eŋ]</td>
</tr>
<tr>
<td>/ē</td>
<td>[ē]</td>
<td>[i̥]</td>
<td>[iŋ]</td>
<td>/i</td>
<td>[i]</td>
<td>[i̥]</td>
<td>[iŋ]</td>
</tr>
<tr>
<td>/ɔ</td>
<td>[ɔ]</td>
<td>[u̥]</td>
<td>[uŋ]</td>
<td>/u</td>
<td>[u]</td>
<td></td>
<td></td>
</tr>
<tr>
<td>/u</td>
<td>[u]</td>
<td>[u̥]</td>
<td>[uŋ]</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

[Note: Evidence from texts written in Brahmi script now suggests that the difference may not be in quantity but in quality, e.g., stressed ē, unstressed ō.]

Examples of nominative singular forms of light- and heavy-stem nouns:

Light stems
βyγ-i <βyγ-y> “god”
βyw- <βyw-y> “god’s”
put- <pwt-y> “Buddha”
marγ- <marγ-y> “bird”
ɔk(r)t- <ɔk(r)t-y> “(was) made”
wirk- <wyrk-y> “wolf”
purn- <pwrn-y> “full”

Heavy stems
βyγ <βyγ-y> “piece of land, garden”
rēz <ryz> “pleasure”
wīnā <wyn> “lute”
rō<s rōz> “copper”
pūt <pwt> “rotted”
māry <marγ-y> “meadow”
mārti <mrti> “man”
pāšmır <ptšmirt> “is (being) counted”
bēndam <bynd(’m) “I bind”
kārθ <kwrθ> “where”

Note: Heavy stems with ir and ur are very rare.
CONSONANTS PHONEMES.

<table>
<thead>
<tr>
<th></th>
<th>Stops</th>
<th>Affricates</th>
<th>Fricatives</th>
<th>Nasals</th>
<th>Continuants</th>
<th>Sibilants</th>
</tr>
</thead>
<tbody>
<tr>
<td>Unvoiced</td>
<td>p, t, k</td>
<td>ċ [ts]</td>
<td>f, θ, x</td>
<td>m, n, [ŋ]</td>
<td>w, y, r</td>
<td>s, š</td>
</tr>
<tr>
<td>Voiced</td>
<td>{b, d, g}</td>
<td>{j}, [dʒ]</td>
<td>β, δ, γ</td>
<td>m, n, [ŋ]</td>
<td>w, y, r</td>
<td>z, ž</td>
</tr>
</tbody>
</table>

In loanwords we also find l and h.

The affricates č and j are pronounced like English ch in child and j in judge.

The fricatives are pronounced as follows: f as in English; θ as English th in thing; x as German ch in Loch or Spanish Spanish (not American Spanish) j in rojo; β, δ, γ are pronounced like Spanish b, d, g after vowels, e.g., robar, nada, haga. The sibilants š and ž are pronounced like English sh in shut and s in pleasure, respectively.

[nỹ], pronounced like English ng in thing, is the phonetic realization of n before k, g, and x. It is not a separate phoneme in Sogdian, only an allophone of /n/.

{b, d, g} and {j} are allophones of /p, t, k, čl/ after the voiced consonants β, δ, γ, m, n, z, ž.

PHONEMES.

We call “phonemes” the smallest units of speech that distinguish meanings. Phonemes are usually determined by establishing “minimal pairs,” for instance, English bad ~ sad, a pair that establishes English /b/ and /s/ as separate phonemes. Phonemes are denoted by writing them between / /. The phoneme is not a sound, merely a linguistic abstraction. When we want to emphasize that we are talking about the actual /b/ and /s/ as separate phonemes. Phonemes are denoted by writing them between / /. The phoneme is not determined by establishing “minimal pairs,” for instance, English

Phonemes may not be distinguished in all positions. Thus, in English we cannot find any minimal pairs distinguished by the presence or absence of voicing in a nasal /m/.

Note that English r is sometimes aspirated [tʰ], sometimes not aspirated [t]. The feature “aspiration” is not, however, distinctive in English or Sogdian, so there is no phonemic opposition /h/ ~ /h’, /p/ ~ /p’, /t/ ~ /t’, /k/ ~ /k’, etc. In this case we say that [p] and [p’] are “allophones” of the phoneme /p/. Aspiration is a distinctive feature in Sanskrit, for instance, where we have minimal pairs such as kara [kara] “hand” ~ khara [k’ara] “donkey.”

Phonemes may not be distinguished in all positions. Thus, in English we cannot find any minimal pairs distinguished by the phoneme sequences /st/ and /sd/. In such cases we say that the phonemic opposition between /t/ and /d/ has been neutralized. Such phonemic neutralization has important consequences for the orthography of Sogdian.

SPECIAL SCRIBAL HABITS.

As the Sogdian alphabets were not created specifically for Sogdian, there is some lack of internal logic in the way letters correspond to sounds. Thus some phonemes are not distinguished in the alphabet (⟨k⟩ = /k/ and /kh/), while others can be written with two different letters. Such optional spellings occur in two situations:

• ⟨č⟩ and ⟨q⟩ both spell k, ⟨t⟩ and ⟨t⟩ both spell t.

• As voiced and unvoiced stops are not distinguished after a voiced consonant, either consonant may be used—for instance, p or b after m, z, and θ; t and d after β, γ, n, z, and š (δ is not found in such combinations); and k and g after n [ŋ], z, and ž. Examples: ⟨ŋb⟩ and ⟨ŋb⟩ mbn = ōʔkāmban “lady, wife.” In the case of nd the most frequent spelling is ⟨nd⟩, less frequently we find ⟨nd⟩, least frequently ⟨nt⟩. Thus, the present participle -andē may be written -ndyy⟩, -ndyy⟩, or -ntyy⟩. Before p and b the opposition between n and m is also neutralized, and either ⟨n⟩ or ⟨m⟩ can be used. Some sound and spelling combinations that occur frequently are the following:
Some letters can be and are frequently written double: <'> (when = ŏ); <ny> and <ww>, especially at the end of words; <ṅṅ> for both ṇ and ñ; <ññ> for n, and <tt> or <tʲt> for t (d).

Final <'> alternates with <-h> (Aramaic hē), and a <-h> can be added after final <'>, <y>, and, occasionally, <w> without affecting the form of the word.

There is no grammatical significance to these alternations. Thus, both βηγι and mártiy can be written with final <γ>, <yy>, or <yh>.

The only grammatical correlation is found in the use of final <h>, which is frequently used with feminine nouns, pronouns, and adjectives, e.g., <wnh> = <wn' >, <xh> = <x', x'' >.

In the grammar and vocabularies in this manual a simplified transliteration system of Manichean Sogdian is used:

- pointed brackets < > are dispensed with;
- letters are written single, not double;
- <k> and <t> are used for <q> and <št>;
- <>' or nothing is used for <h> (<βγ> not <βγh>, <mrty> not <mrtyh>, etc.).

EXERCISES 1

1. Read and transcribe the following words:

2. Suggest spellings for the following transcribed words, and write them in Manichean script:

GLOSSARY 1

Learn the following words by heart:

- 'ykwn ayykōn: eternally
- 'zyn dēn: parable, story
- 'sp xsp: horse
- βr't βrnr: brother
- ŧhr - ŧhrθor - θθr: to give, given
- ñs' ñsr: ten
- ŧyn' ḏnr (or ðnær): dinar
- fryšt fhrēšē: angel
- γ'swk γadhūk: throne
- mrty marṭiy: man
- myô mēθ: day
- myô, m'yô mēð: thus
- m't mār fem.: mother

nwkr nākar: now
nwr nūr: today
ptr pātew: father
pts' patsēr: again, once more
ptvĬş- ptvĬšt ptvĬšē: to hear, heard
š'twxs šāux: glad, happy
wn un (wun) fem.: tree
x'n' xāndā fem.: house
xwt' w xūtāw: lord, king
zrw'βy zrwā-βγr: God Zurwān, the Father of Greatness
zyrn zērn: gold
### TABLE OF CORRESPONDENCES SOUND – SPELLING

<table>
<thead>
<tr>
<th>Transcription</th>
<th>Spelling</th>
<th>Transcription</th>
<th>Spelling</th>
</tr>
</thead>
<tbody>
<tr>
<td>a (ə, i)</td>
<td>’, nothing</td>
<td>l</td>
<td>l</td>
</tr>
<tr>
<td>ā</td>
<td>’ . ’</td>
<td>m</td>
<td>m n</td>
</tr>
<tr>
<td>-ā</td>
<td>-h -’</td>
<td>n</td>
<td>n</td>
</tr>
<tr>
<td>b in mb</td>
<td>b p</td>
<td>ō</td>
<td>w ww</td>
</tr>
<tr>
<td>β</td>
<td>β</td>
<td>p</td>
<td>p</td>
</tr>
<tr>
<td>ĉ</td>
<td>c</td>
<td>r</td>
<td>r</td>
</tr>
<tr>
<td>d in nd, ndt</td>
<td>d dt t</td>
<td>s</td>
<td>s</td>
</tr>
<tr>
<td>ď</td>
<td>ď</td>
<td>ŝ</td>
<td>ŝ</td>
</tr>
<tr>
<td>ě</td>
<td>ř yy</td>
<td>t</td>
<td>t ř t ĭ</td>
</tr>
<tr>
<td>f</td>
<td>ľ</td>
<td>ř</td>
<td>ř δ</td>
</tr>
<tr>
<td>g in ng</td>
<td>g k q</td>
<td>ŭ</td>
<td>w ww</td>
</tr>
<tr>
<td>γ</td>
<td>γ</td>
<td>w</td>
<td>w</td>
</tr>
<tr>
<td>h (rare)</td>
<td>h (ḥ)</td>
<td>x</td>
<td>x</td>
</tr>
<tr>
<td>Ĩ</td>
<td>Ĩ yy</td>
<td>y</td>
<td>y</td>
</tr>
<tr>
<td>Ĩ</td>
<td>Ĩ</td>
<td>z</td>
<td>z</td>
</tr>
<tr>
<td>J in nj</td>
<td>c</td>
<td>ż</td>
<td>j</td>
</tr>
<tr>
<td>k</td>
<td>k q</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>