

Universal Multiple-Octet Coded Character Set
International Organization for Standardization
Organisation Internationale de Normalisation
Международная организация по стандартизации

Doc Type: Working Group Document**Title: Preliminary proposal to encode the Book Pahlavi script in the BMP of the UCS****Source: Michael Everson, Roozbeh Pournader, and Desmond Durkin-Meisterernst****Status: Individual Contribution****Action: For consideration by JTC1/SC2/WG2 and UTC****Replaces: N2556****Date: 2007-07-30**

1. Introduction. “Pahlavi” is a term used in two senses. As a term for language, it means Zoroastrian or Sasanian Middle Persian. As a term for writing systems, the word is used to describe the scripts used to write religious and secular Sasanian Middle Persian and closely similar material, such as epigraphic Parthian and Middle Persian. Three “Pahlavi” scripts are distinguished: Inscriptional Pahlavi, Psalter Pahlavi, and Book Pahlavi. All of these derive from Imperial Aramaic.

Book Pahlavi is an alphabetic script with an incomplete representation of vowels. The original Imperial Aramaic script’s development is marked by a progressive loss of differentiation of many letters. Thus the Imperial Aramaic letters 𐤱 WAW, 𐤬 NUN, 𐤣 AYIN, and 𐤤 RESH coalesced into one simple vertical stroke, in Book Pahlavi written 𐬀. (In Inscriptional Pahlavi only 𐤱 WAW, 𐤬 NUN, and 𐤤 RESH coalesce into 𐤢.) The Imperial Aramaic letters 𐤀 ALEPH and 𐤁 HET coalesced into Book Pahlavi 𐬁; the Imperial Aramaic letters 𐤂 GIMEL, 𐤃 DALETH, and 𐤄 YODH coalesced into Book Pahlavi 𐬂; and (less importantly) the Imperial Aramaic letters 𐤅 MEM and 𐤆 QOPH coalesced into Book Pahlavi 𐬃.

In addition to the confusion caused by this lack of differentiation, Book Pahlavi made use of many ligatures, and often their readings are quite ambiguous so that a number of readings may be possible for a given word-form. The deterioration and resulting ambiguity has had important consequences both for the scribes who copied the texts and for modern editors trying to edit them. The scribes, unable to read with certainty the manuscripts they were copying, introduced further confusion into already difficult texts. Modern editors often cannot arrive at what the text originally contained. Each editor can only offer his interpretation of what is written and though theoretically only one interpretation should be correct no editor can prove his interpretation to the exclusion of others based on the written forms alone. Some parts of Pahlavi literature are quite simply irrecoverable despite being transmitted in manuscripts.

The response to this ambiguity was the development of the Avestan script. Avestan used modified letterforms to distinguish between signs which in Pahlavi had fallen together. Indeed, while Pahlavi words are commonly used within Avestan texts, it was often the case that the scribes had no idea what the values of the words were, because so many of the letters’ shapes were identical. The notion of *archigrapheme* has been found to be useful in proposing several of the Pahlavi extension characters: it is often quite clear that the scribes themselves were writing “a bowl and an ear” even though no such letter exists in Pahlavi. These archigraphemes are proposed to be encoded in order to enable scholars to represent the same kinds of ambiguities which the scribes themselves were writing. This is not the same as encoding palaeographical variants.

2. Orthography. Book Pahlavi is a mixed orthographic system, making extensive use of Aramaic spellings to represent the equivalent Middle Persian words. This heterographic principle is similar to the use of Chinese characters together with Japanese syllabic signs to write Japanese or the use of Sumerian

words in Akkadian texts. For instance, 𐭠𐭣𐭥 MLK' writes Middle Persian *šāh* 'king', 𐭠𐭣𐭥 YMLLWN writes Middle Persian *gōw* 'say!' (the heterograms are transliterated with capital letters). As Japanese does with its Han characters, phonetic complements are added to the heterograms to indicate a specific ending: 𐭠𐭣𐭥 MLK"n for *šāhān* 'kings' 𐭠𐭣𐭥 and YMLLWNyt for *gōwēd* 'he says'.

Book Pahlavi orthography is also archaizing in that it uses historical spellings. For example, 𐭠𐭣𐭥 *l'd* spells *rāy* 'for', 𐭠𐭣𐭥 *ywk* spells *ēk* 'one', 𐭠𐭣𐭥 *y'wl* spells *jār* 'time'; this means that some letters seem to shift values ($d > y$; $y > j$). In fact, the sound changes are simply not expressed by the letters which write the underlying and, in most cases, historically earlier forms.

3. Processing. Book Pahlavi is written right-to-left. Usually words are separated with spaces; sometimes avestan word separator is used to separate word. Book Pahlavi script has fully-developed joining behaviour. The table below shows the joining forms.

[To be supplied. This table may prove to be extremely complex. There is a real question as to whether there is joining behaviour in Book Pahlavi as in other RTL scripts, or if it is all a massive set of ligatures.]

4. Book Pahlavi punctuation. Punctuation as in Avestan is used. There may be some additional triple dots; more research is needed. Two logograms are used: an ABBREVIATION TAA 'thus' and the inverted name of Ahreman, 'the Evil Spirit'.

5. Book Pahlavi numbers. Book Pahlavi has its own numbers, which have right-to-left directionality. Numbers are built up out of 1, 2, 3, 4, 10, 20, 40, 60, 80, and 100. The following is a list of numbers attested in Book Pahlavi. The third column is displayed in visual order.

[To be supplied.]

6. Names and ordering. The names used for the Book Pahlavi characters are based on their Imperial Aramaic analogues. The order of the characters in the code charts is their alphabetical order. The historical characters GIMEL, DALETH, and YODH fell together into a single character, named here GIMEL-DALETH-YODH. The same thing happened to ALEPH-HET, WAW-NUN-AYIN-RESH and MEM-QOPH respectively.

7. Consequences for encoding. Consequences for encoding Pahlavi arise from the first feature in particular. The progressive loss of differentiation of many letters means that much Pahlavi writing is ambiguous and some simply unintelligible. Though the Zoroastrian tradition, the editor's experience and reference to Manichaean Middle Persian can help resolve many difficulties there are many cases where the same original text is read differently by as many editors as have worked on the text and all the readings are and remain plausible interpretations of what is written in the manuscript. There are also some cases where an editor despairs of reading a passage and has to reproduce the original shapes without being able to venture a reading. Later editors may find the solution but there will always be unresolved and irresolvable cases.

8. Unicode Character Properties. Character properties are proposed here. With regard to the combining classes of the vowel signs, normalization should weight the combining marks in five groups:

```
10BA0;BOOK PAHLAVI LETTER ALEPH-HET;Lo;0;R;;;N;;;;;
10BA1;BOOK PAHLAVI LETTER BETH;Lo;0;R;;;N;;;;;
10BA2;BOOK PAHLAVI LETTER GIMEL-DALETH-YODH;Lo;0;R;;;N;;;;;
10BA3;BOOK PAHLAVI LETTER HE;Lo;0;R;;;N;;;;;
10BA4;BOOK PAHLAVI LETTER WAW-NUN-AYIN-RESH;Lo;0;R;;;N;;;;;
10BA5;BOOK PAHLAVI LETTER ZAYIN;Lo;0;R;;;N;;;;;
10BA6;BOOK PAHLAVI LETTER KAPH;Lo;0;R;;;N;;;;;
10BA7;BOOK PAHLAVI LETTER GHAPH;Lo;0;R;;;N;;;;;
10BA8;BOOK PAHLAVI LETTER LAMEDH;Lo;0;R;;;N;;;;;
```

10BA9;BOOK PAHLAVI LETTER LHAMEDH;Lo;0;R;;;N;;;;;
 10BAA;BOOK PAHLAVI LETTER MEM-QOPH;Lo;0;R;;;N;;;;;
 10BAB;BOOK PAHLAVI LETTER SAMEKH;Lo;0;R;;;N;;;;;
 10BAC;BOOK PAHLAVI LETTER PE;Lo;0;R;;;N;;;;;
 10BAD;BOOK PAHLAVI LETTER SADHE;Lo;0;R;;;N;;;;;
 10BAE;BOOK PAHLAVI LETTER SHIN;Lo;0;R;;;N;;;;;
 10BAF;BOOK PAHLAVI LETTER TAW;Lo;0;R;;;N;;;;;
 10BB0;BOOK PAHLAVI ARCHIGRAPHME EAR;Lo;0;R;;;N;;;;;
 10BB1;BOOK PAHLAVI ARCHIGRAPHME ELBOW;Lo;0;R;;;N;;;;;
 10BB2;BOOK PAHLAVI ARCHIGRAPHME BELLY;Lo;0;R;;;N;;;;;
 10BBA;BOOK PAHLAVI COMBINING GIMEL;Mn;230;NSM;;;N;;;;;
 10BBB;BOOK PAHLAVI COMBINING DALETH;Mn;230;NSM;;;N;;;;;
 10BBC;BOOK PAHLAVI COMBINING YODH;Mn;220;NSM;;;N;;;;;
 10BBD;BOOK PAHLAVI COMBINING SAMEKH;Mn;220;NSM;;;N;;;;;
 10BBE;BOOK PAHLAVI COMBINING SHIN;Mn;230;NSM;;;N;;;;;
 10BBF;BOOK PAHLAVI KASHIDA;Lm;0;AL;;;N;;;;;
 10BC0;BOOK PAHLAVI ABBREVIATION TAA;Lo;0;R;;;N;;;;;
 10BC1;BOOK PAHLAVI LOGOGRAM TURNED AHREMAN;Lo;0;R;;;N;;;;;
 10BC5;BOOK PAHLAVI NUMBER ONE;No;0;R;;;1;N;;;;;
 10BC6;BOOK PAHLAVI NUMBER TWO;No;0;R;;;2;N;;;;;
 10BC7;BOOK PAHLAVI NUMBER THREE;No;0;R;;;3;N;;;;;
 10BC8;BOOK PAHLAVI NUMBER FOUR;No;0;R;;;4;N;;;;;
 10BC9;BOOK PAHLAVI NUMBER TEN;No;0;R;;;10;N;;;;;
 10BCA;BOOK PAHLAVI NUMBER TWENTY;No;0;R;;;20;N;;;;;
 10BCB;BOOK PAHLAVI NUMBER FORTY;No;0;R;;;40;N;;;;;
 10BCC;BOOK PAHLAVI NUMBER SIXTY;No;0;R;;;60;N;;;;;
 10BCD;BOOK PAHLAVI NUMBER EIGHTY;No;0;R;;;80;N;;;;;
 10BCE;BOOK PAHLAVI NUMBER ONE HUNDRED;No;0;R;;;100;N;;;;;
 10BCF;BOOK PAHLAVI NUMBER ONE THOUSAND;No;0;R;;;1000;N;;;;;

9. Bibliography.

- Akbarzādeh, Dāriyūš. 2002 (1381 AP). *Katibe-hā-ye Pahlavi-ye aškāni (Pārti) = Parthian inscriptions*. Vol. II. Tehran: Pazineh Press. ISBN 964-5722-74-8
- Ballhorn, Freidrich. 1864. *Alphabete orientalischer und occidentalischer Sprachen*. Neunte vermehrte Auflage. Leipzig: F. A. Brockhaus.
- de Harlex, C. 1880. *Manuel du pehlevi des livres religieux et historiques de la perse. Grammaire, anthologie lexique avec des notes, un fac-simile de manuscrit les alphabets et un spécimen des légendes des sceaux et monnaies*. Paris: Maisonneuve et Cie.
- Faulmann, Carl. 1990 (1880). *Das Buch der Schrift*. Frankfurt am Main: Eichborn. ISBN 3-8218-1720-8
- Geldner, Karl F. 1880. *Avesta: the sacred books of the Parsis*. Stuttgart: W. Kohlhammer. Reprinted in 2003 with an introduction in Persian by Dr Jaleh Amouzgar.
- Kōno Rokurō, Chino Eiichi, & Nishida Tatsuo. 2001. *The Sanseido Encyclopaedia of Linguistics. Volume 7: Scripts and Writing Systems of the World [Gengogaku dai ziten (bekkan) sekai mozi ziten]*. Tokyo: Sanseido Press. ISBN 4-385-15177-6
- MacKenzie, D. N. 1971. *A concise Pahlavi dictionary*. London: Oxford University Press.
- Nyberg, Henrik Samuel. 1964 *A manual of Pahlavi*. Wiedbaden: Otto Harrassowitz. Reprinted 2003 Tehran: Asatir. ISBN 964-331-131-7, 964-331-132-5
- Reichsdruckerei. 1951. *Alphabete und Schriftzeichen des Morgen- und Abendlandes, zum allgemeinen Gebrauch mit besonderer Berücksichtigung des Buchgewerbes. Unter Mitwirkung von Fachgelehrten zusammengestellt in der Reichsdruckerei*. Berlin: Staatsdruckerei.
- Skjærvø, P. Oktor. 1996. "Aramaic scripts for Iranian languages" in *The World's Writing Systems*, ed. Peter T. Daniels & William Bright. New York; Oxford: Oxford University Press. ISBN 0-19-507993-0

10. Acknowledgements

This project was made possible in part by a grant from the U.S. National Endowment for the Humanities, which funded the Universal Scripts Project (part of the Script Encoding Initiative at UC Berkeley). Any views, findings, conclusions or recommendations expressed in this publication do not necessarily reflect those of the National Endowment of the Humanities.

Figures

THE IRANIAN ALPHABETS.

	ARAMEAN.		PEHLEVI.					INDO- BACTRIAN.	ARMENIAN. (Reversed.)	GEORGIAN. (Reversed.)
			ARSACIDAN.		SASSANIAN.		PARSI.			
	SATRA- PIES & EGYPT.	PAL- MYRA.	Coins & Gems.	Haji- abad B.	Haji- abad A.	Coins.	MS.	KAPUR- DI-GIRI.	Oldest MSS.	Oldest MSS.
	Sec. iv. & iii. B.C.	Sec. ii. A.D.	Sec. i. & ii. A.D.	Sec. iii. A.D.	Sec. iii. A.D.	Sec. iv. to vi. A.D.	Modern	Sec. iii. B.C.	Sec. ix. A.D.	Sec. x. A.D.
א	𐤀	𐤁	𐭀𐭁	𐭀𐭂	𐭀	𐭀	𐭀	𐭀	𐭀	𐭀
ב	𐤁	𐤂	𐭀	𐭀	𐭀	𐭀	𐭀	𐭀	𐭀	𐭀
ג	𐤂	𐤃	𐭀𐭂	𐭀𐭃	𐭀𐭃	𐭀𐭃	𐭀	𐭀	𐭀	𐭀
ד	𐤃	𐤄	𐭀	𐭀	𐭀	𐭀	𐭀	𐭀	𐭀	𐭀
ה	𐤄	𐤅	𐭀𐭄	𐭀𐭅	𐭀𐭅	𐭀𐭅	𐭀	𐭀	𐭀	𐭀
ו	𐤅	𐤆	𐭀𐭅	𐭀𐭆	𐭀𐭆	𐭀𐭆	𐭀	𐭀	𐭀	𐭀
ז	𐤆	𐤇	𐭀	𐭀	𐭀	𐭀	𐭀	𐭀	𐭀	𐭀
ח	𐤇	𐤈	𐭀	𐭀	𐭀	𐭀	𐭀	𐭀	𐭀	𐭀
ט	𐤈	𐤉	𐭀	𐭀	𐭀	𐭀	𐭀	𐭀	𐭀	𐭀
י	𐤉	𐤊	𐭀𐭉	𐭀𐭉	𐭀𐭉	𐭀𐭉	𐭀	𐭀	𐭀	𐭀
כ	𐤊	𐤋	𐭀𐭉	𐭀𐭉	𐭀𐭉	𐭀𐭉	𐭀	𐭀	𐭀	𐭀
ל	𐤋	𐤌	𐭀𐭉	𐭀𐭉	𐭀𐭉	𐭀𐭉	𐭀	𐭀	𐭀	𐭀
מ	𐤌	𐤍	𐭀𐭉	𐭀𐭉	𐭀𐭉	𐭀𐭉	𐭀	𐭀	𐭀	𐭀
נ	𐤍	𐤎	𐭀𐭉	𐭀𐭉	𐭀𐭉	𐭀𐭉	𐭀	𐭀	𐭀	𐭀
ס	𐤎	𐤏	𐭀𐭉	𐭀𐭉	𐭀𐭉	𐭀𐭉	𐭀	𐭀	𐭀	𐭀
ע	𐤏	𐤐	𐭀𐭉	𐭀𐭉	𐭀𐭉	𐭀𐭉	𐭀	𐭀	𐭀	𐭀
פ	𐤐	𐤑	𐭀𐭉	𐭀𐭉	𐭀𐭉	𐭀𐭉	𐭀	𐭀	𐭀	𐭀
צ	𐤑	𐤒	𐭀𐭉	𐭀𐭉	𐭀𐭉	𐭀𐭉	𐭀	𐭀	𐭀	𐭀
ק	𐤒	𐤓	𐭀𐭉	𐭀𐭉	𐭀𐭉	𐭀𐭉	𐭀	𐭀	𐭀	𐭀
ר	𐤓	𐤔	𐭀𐭉	𐭀𐭉	𐭀𐭉	𐭀𐭉	𐭀	𐭀	𐭀	𐭀
ש	𐤔	𐤕	𐭀𐭉	𐭀𐭉	𐭀𐭉	𐭀𐭉	𐭀	𐭀	𐭀	𐭀
ת	𐤕	𐤖	𐭀𐭉	𐭀𐭉	𐭀𐭉	𐭀𐭉	𐭀	𐭀	𐭀	𐭀

(236)

Figure 1. Table of Iranian alphabets, from Taylor 1883. The Arsacidan columns (III, IV) show Parthian script, and the Sassanian columns (V, VI) show Inscriptional Pahlavi script. The Parsi column (VII) shows **Book Pahlavi**. Column I shows Imperial Aramaic; column II shows Palmyrene; column VIII shows Brahmi; column IX shows Armenian; and column X shows Georgian.

This table is given for historical interest. It stems from a period when the inscriptions were known but could not yet be properly read—and that did not happen until the 1920s. Taylor's table is therefore inaccurate: under Parthian DALETH he has the letter TETH, the form of LAMEDH is wrong, and he gives no SADHE. Under Middle Persian, TETH is missing and TAW has the wrong shape.

Alphabets.

Hebrew	Kesā'	Aramaic papyri	Palmyr. inscrip- tions	Nabat. inscrip- tions, Sinai	Parthian inscrip- tions	Persian inscrip- tions	Psalter	Book Pahlavi	Trans- liter- ation
א	𐎀	𐤀	𐤀	𐤀	𐭀	𐭀	𐭀	𐭀	'
ב	𐎁	𐤁	𐤁	𐤁, 𐤂	𐭁	𐭁	𐭁	𐭁	b
ג	𐎂	𐤂	𐤂	𐤂, 𐤃	𐭂	𐭂	𐭂	𐭂 (ḡ)	g
ד	𐎃	𐤃	𐤃	𐤃, 𐤄	𐭃	𐭃	𐭃	𐭃 (ḥ) (ḡ-)	d
ה	𐎄	𐤄	𐤄	𐤄, 𐤅	𐭄	𐭄	𐭄	𐭄 (ḥ) (ḡ-)	-h
ו	𐎅	𐤅	𐤅	𐤅, 𐤆	𐭅	𐭅	𐭅	𐭅	w
ז	𐎆	𐤆	𐤆	𐤆, 𐤇	𐭆	𐭆	𐭆	𐭆	z
ח	𐎇	𐤇	𐤇	𐤇, 𐤈	𐭇	𐭇	𐭇	𐭇	h, ḥ
ט	𐎈	𐤈	𐤈	𐤈, 𐤉	𐭈	𐭈	𐭈	𐭈	ṭ
י	𐎉	𐤉	𐤉	𐤉, 𐤊	𐭉	𐭉	𐭉	𐭉 (ḡ, ḡ-)	y (j-)
כ. 𐎁	𐎁	𐤁	𐤁	𐤁, 𐎁	𐭁	𐭁	𐭁	𐭁	k
ל	𐎂	𐤂	𐤂	𐤂, 𐎂	𐭂	𐭂	𐭂	𐭂 (b)	l
מ. 𐎁	𐎃	𐤃	𐤃	𐤃, 𐎁, 𐎂	𐭃	𐭃	𐭃	𐭃	m
נ. 𐎁	𐎄	𐤄	𐤄	𐤄, 𐎁	𐭄	𐭄	𐭄	𐭄	n
ס	𐎅	𐤅	𐤅	𐤅	𐭅	𐭅	𐭅	𐭅, -ṭ	s
ע	𐎆	𐤆	𐤆	𐤆, 𐎁	𐭆	𐭆	𐭆	𐭆	c
פ. 𐎁	𐎇	𐤇	𐤇	𐤇, 𐎁	𐭇	𐭇	𐭇	𐭇	p
צ. 𐎁	𐎈	𐤈	𐤈	𐤈, 𐎁	𐭈	𐭈	𐭈	𐭈	c, s
ק	𐎉	𐤉	𐤉	𐤉, 𐎁	𐭉	𐭉	𐭉	𐭉	q
ר	𐎁	𐤁	𐤁	𐤁, 𐎁, 𐎂	𐭁	𐭁	𐭁	𐭁	r
ש	𐎂	𐤂	𐤂	𐤂, 𐎁, 𐎂	𐭂	𐭂	𐭂	𐭂	š
ת	𐎃	𐤃	𐤃	𐤃, 𐎁	𐭃	𐭃	𐭃, 𐭄	𐭃	t

Figure 2. Table of Iranian alphabets, from from Nyberg 1964.

Book Pahlavi is given toward the right of the table. The Parthian inscriptions column refers to Parthian, the Persian inscriptions column refers to Inscriptional Pahlavi. Psalter Pahlavi is also shown. Nyberg's table is slightly idiosyncratic and the shape of Parthian ALEPH is wrong.

TABLE 48.1: *Main West Iranian Scripts Developed from Aramaic*

<i>Aramaic</i>	<i>Early Parthian (Nisa)</i>	<i>Parthian inscr's</i>	<i>Principal Phonetic Values (Parthian)</i>	<i>Middle Persian inscr's</i>	<i>Psalter</i>	<i>Early Cursive Pahlavi</i>	<i>Book Pahlavi</i>	<i>Principal Phonetic Values (MPers.)</i>
ʾ	𐭪	𐭪	a, ā	𐭪	𐭪	𐭪	𐭪	a, ā
b	𐭫	𐭫	b, w	𐭫	𐭫	𐭫	𐭫	b, w
g	𐭬	𐭬	g, γ	𐭬	𐭬		𐭬 𐭬	g, γ
d	𐭭 = ʿ, r	𐭭	d, δ	𐭭	𐭭	𐭭	𐭭 𐭭	d, y
h (h)	𐭮	𐭮		𐭮	𐭮		𐭮	
w	𐭯	𐭯	w, ō, ū	2 = ʿ, r	1 = ʿ, r	1	1	w, ō, ū
z	𐭰	𐭰	z, ž	𐭰	𐭰		𐭰	z
h (h)	𐭱	𐭱	h, x	𐭱	𐭱	𐭱	𐭱	h, x
ʔ	𐭲	𐭲		𐭲				
y	𐭳	𐭳	y, ē, ī	𐭳	𐭳	𐭳	𐭳 𐭳	y, ē, ī, j
k	𐭴	𐭴	k, g	𐭴	𐭴	𐭴	𐭴	k, g
l (δ)	𐭵	𐭵	l	𐭵	𐭵	𐭵	𐭵	l, r
m	𐭶	𐭶	m	𐭶	𐭶	𐭶	𐭶	m
n	𐭷	𐭷	n	𐭷	𐭷	𐭷	𐭷	n
s	𐭸	𐭸	s	𐭸	𐭸	𐭸	𐭸 𐭸	s, h
ʿ	𐭹 = d, r	𐭹 = r		2 = w, r	1 = w, r	1	1	Ø
p	𐭺	𐭺	p, b	𐭺	𐭺	𐭺 𐭺	𐭺	p, b, f
š (c)		𐭻	č	𐭻	𐭻	𐭻	𐭻 𐭻	č, j, z
q	𐭼	𐭼		𐭼	𐭼	𐭼	𐭼	
r	𐭽 = d, ʿ	𐭽 = ʿ	r	2 = w, ʿ	1 = w, ʿ	1	1	r
š	𐭾	𐭾	š, ž	𐭾	𐭾	𐭾	𐭾	š
t	𐭿	𐭿	t, d	𐭿	𐭿	𐭿	𐭿	t, d

Figure 3. Table of Iranian alphabets, from from Skjærvø 1996.

TABLE I					
Imperial Aramaic		Inscriptions		Pahlavi	
		Parthian	Persian	Psalter	Book
' ('āleṗ)*	𐤀	𐭀	𐭀	𐭀	𐭀
b (bēṭ)	𐤁	𐭁	𐭁	𐭁	𐭁
g (gimel)	𐤂	𐭂	𐭂	𐭂	𐭂
d (dāleṭ)	𐤃	𐭃	𐭃	𐭃	𐭃 = g
h (hē)	𐤄	𐭄	𐭄	𐭄	𐭄
w (wāw)	𐤅	𐭅	𐭅	𐭅	𐭅
z (zayin)	𐤆	𐭆	𐭆	𐭆	𐭆
ḥ (ḥēṭ)	𐤇	𐭇	𐭇	𐭇	𐭇
t (tēṭ)	𐤈	𐭈	𐭈	𐭈	𐭈
y (yōḏ)	𐤉	𐭉	𐭉	𐭉	𐭉 = g
k (kaṗ)	𐤊	𐭊	𐭊	𐭊	𐭊
l (lāmed)	𐤋	𐭋	𐭋	𐭋	𐭋
m (mēm)	𐤌	𐭌	𐭌	𐭌	𐭌
n (nūn)	𐤍	𐭍	𐭍	𐭍	𐭍 = w
s (sāmēḵ)	𐤎	𐭎	𐭎	𐭎	𐭎
' ('ayin)	𐤏	𐭏	= w	= w	= w
p (pē)	𐤐	𐭐	𐭐	𐭐	𐭐
ṣ (ṣāḏḏē)	𐤑	𐭑	𐭑	𐭑	𐭑
q (qōṗ)	𐤒	𐭒	= m	= m	= m
r (rēṣ)	𐤓	=	= w	= w	= w
ś ś (ś/šīn)	𐤔	𐭔	𐭔	𐭔	𐭔
t (tāw)	𐤕	𐭕	𐭕	𐭕	𐭕

* The under- and over-lining of letters in Semitic transcription indicates a fricative pronunciation, *āleph*, *bēth*, etc.

Figure 4. Table showing Imperial Aramaic, Inscriptional Parthian, Inscriptional Pahlavi, and Book Pahlavi, from MacKenzie 1971.

جدول الفبای پهلوی

حرف نویسی	زبور پهلوی	پهلوی کتابی	پهلوی کتیبه ای
Ⲁ A	𐭀	𐭀	𐭀
b	𐭁	𐭁	𐭁
g	𐭂	𐭂/𐭃	𐭂
d	𐭃	𐭃/𐭄	𐭃
E	𐭄	𐭄	𐭄
w	𐭅	𐭅	𐭅
z	𐭆	𐭆	𐭆
h	𐭇	𐭇	𐭇
t	𐭈	𐭈	𐭈
y	𐭉/𐭊	𐭉/𐭊	𐭉
k	𐭊	𐭊	𐭊
l	𐭋	𐭋	𐭋
m	𐭌	𐭌	𐭌
n	𐭍	𐭍	𐭍
s	𐭎	𐭎/𐭏	𐭎
O	𐭏	𐭏	𐭏
P	𐭐	𐭐	𐭐
c	𐭑	𐭑	𐭑
Q	𐭒	𐭒	𐭒
r	𐭓	𐭓	𐭓
s	𐭔	𐭔	𐭔
t	𐭕	𐭕	𐭕

Figure 5. Table of Iranian alphabets, showing Psalter Pahlavi on the left, Book Pahlavi in the centre, and Inscriptional Pahlavi on the right, from Akbarzadeh 2002.

PEHLEWI

Zeichen	Lautwert	Zeichen	Lautwert
𐭌	a, ā (besser o, ō), h, x	𐭌𐭎	č (z)
𐭎	i, ē, y, d (ð), g (γ)	𐭎	t (ð)
𐭏	i, ē, y	𐭏	p, f (β)
𐭐	dž	𐭐	m, q
𐭑	g (γ)	𐭑	r, l
𐭒	d (ð)	𐭒	l
𐭓	Schluß - ē, b (β)	𐭓	s
𐭔	{u, ū, o, ō, v, n, r, Ain	𐭔𐭕	š
𐭕	k (γ)	𐭕	z
𐭖	γ	𐭖	h
Ligaturen			
𐭗	𐭗, 𐭗, 𐭗, 𐭗	𐭗	𐭗, 𐭗
𐭘	𐭘	𐭘	𐭘
𐭙	𐭙, 𐭙, 𐭙, 𐭙	𐭙	𐭙, 𐭙
𐭚	𐭚	𐭚	𐭚, 𐭚
𐭛	𐭛, 𐭛	𐭛	𐭛, 𐭛
𐭜	𐭜, 𐭜, 𐭜, 𐭜	𐭜	𐭜
𐭝	𐭝	𐭝	𐭝, 𐭝
𐭞	𐭞	𐭞	𐭞
𐭟	𐭟	𐭟	𐭟, 𐭟
𐭠	𐭠	𐭠	𐭠
𐭡	𐭡	𐭡	𐭡, 𐭡
𐭢	𐭢	𐭢	𐭢
𐭣	𐭣	𐭣	𐭣, 𐭣
𐭤	𐭤	𐭤	𐭤
𐭥	𐭥	𐭥	𐭥, 𐭥
𐭦	𐭦	𐭦	𐭦
𐭧	𐭧	𐭧	𐭧, 𐭧
𐭨	𐭨	𐭨	𐭨
𐭩	𐭩	𐭩	𐭩, 𐭩
𐭪	𐭪	𐭪	𐭪
𐭫	𐭫	𐭫	𐭫, 𐭫
𐭬	𐭬	𐭬	𐭬
𐭭	𐭭	𐭭	𐭭, 𐭭
𐭮	𐭮	𐭮	𐭮
𐭯	𐭯	𐭯	𐭯, 𐭯
𐭰	𐭰	𐭰	𐭰
𐭱	𐭱	𐭱	𐭱, 𐭱
𐭲	𐭲	𐭲	𐭲
𐭳	𐭳	𐭳	𐭳, 𐭳
𐭴	𐭴	𐭴	𐭴
𐭵	𐭵	𐭵	𐭵, 𐭵
𐭶	𐭶	𐭶	𐭶
𐭷	𐭷	𐭷	𐭷, 𐭷
𐭸	𐭸	𐭸	𐭸
𐭹	𐭹	𐭹	𐭹, 𐭹
𐭺	𐭺	𐭺	𐭺
𐭻	𐭻	𐭻	𐭻, 𐭻
𐭼	𐭼	𐭼	𐭼
𐭽	𐭽	𐭽	𐭽, 𐭽
𐭾	𐭾	𐭾	𐭾
𐭿	𐭿	𐭿	𐭿, 𐭿
𐮀	𐮀	𐮀	𐮀
𐮁	𐮁	𐮁	𐮁, 𐮁
𐮂	𐮂	𐮂	𐮂
𐮃	𐮃	𐮃	𐮃, 𐮃
𐮄	𐮄	𐮄	𐮄
𐮅	𐮅	𐮅	𐮅, 𐮅
𐮆	𐮆	𐮆	𐮆
𐮇	𐮇	𐮇	𐮇, 𐮇
𐮈	𐮈	𐮈	𐮈
𐮉	𐮉	𐮉	𐮉, 𐮉
𐮊	𐮊	𐮊	𐮊
𐮋	𐮋	𐮋	𐮋, 𐮋
𐮌	𐮌	𐮌	𐮌
𐮍	𐮍	𐮍	𐮍, 𐮍
𐮎	𐮎	𐮎	𐮎
𐮏	𐮏	𐮏	𐮏, 𐮏
𐮐	𐮐	𐮐	𐮐
𐮑	𐮑	𐮑	𐮑, 𐮑
𐮒	𐮒	𐮒	𐮒
𐮓	𐮓	𐮓	𐮓, 𐮓
𐮔	𐮔	𐮔	𐮔
𐮕	𐮕	𐮕	𐮕, 𐮕
𐮖	𐮖	𐮖	𐮖
𐮗	𐮗	𐮗	𐮗, 𐮗
𐮘	𐮘	𐮘	𐮘
𐮙	𐮙	𐮙	𐮙, 𐮙
𐮚	𐮚	𐮚	𐮚
𐮛	𐮛	𐮛	𐮛, 𐮛
𐮜	𐮜	𐮜	𐮜
𐮝	𐮝	𐮝	𐮝, 𐮝
𐮞	𐮞	𐮞	𐮞
𐮟	𐮟	𐮟	𐮟, 𐮟
𐮠	𐮠	𐮠	𐮠
𐮡	𐮡	𐮡	𐮡, 𐮡
𐮢	𐮢	𐮢	𐮢
𐮣	𐮣	𐮣	𐮣, 𐮣
𐮤	𐮤	𐮤	𐮤
𐮥	𐮥	𐮥	𐮥, 𐮥
𐮦	𐮦	𐮦	𐮦
𐮧	𐮧	𐮧	𐮧, 𐮧
𐮨	𐮨	𐮨	𐮨
𐮩	𐮩	𐮩	𐮩, 𐮩
𐮪	𐮪	𐮪	𐮪
𐮫	𐮫	𐮫	𐮫, 𐮫
𐮬	𐮬	𐮬	𐮬
𐮭	𐮭	𐮭	𐮭, 𐮭
𐮮	𐮮	𐮮	𐮮
𐮯	𐮯	𐮯	𐮯, 𐮯
𐮰	𐮰	𐮰	𐮰
𐮱	𐮱	𐮱	𐮱, 𐮱
𐮲	𐮲	𐮲	𐮲
𐮳	𐮳	𐮳	𐮳, 𐮳
𐮴	𐮴	𐮴	𐮴
𐮵	𐮵	𐮵	𐮵, 𐮵
𐮶	𐮶	𐮶	𐮶
𐮷	𐮷	𐮷	𐮷, 𐮷
𐮸	𐮸	𐮸	𐮸
𐮹	𐮹	𐮹	𐮹, 𐮹
𐮺	𐮺	𐮺	𐮺
𐮻	𐮻	𐮻	𐮻, 𐮻
𐮼	𐮼	𐮼	𐮼
𐮽	𐮽	𐮽	𐮽, 𐮽
𐮾	𐮾	𐮾	𐮾
𐮿	𐮿	𐮿	𐮿, 𐮿
𐯀	𐯀	𐯀	𐯀
𐯁	𐯁	𐯁	𐯁, 𐯁
𐯂	𐯂	𐯂	𐯂
𐯃	𐯃	𐯃	𐯃, 𐯃
𐯄	𐯄	𐯄	𐯄
𐯅	𐯅	𐯅	𐯅, 𐯅
𐯆	𐯆	𐯆	𐯆
𐯇	𐯇	𐯇	𐯇, 𐯇
𐯈	𐯈	𐯈	𐯈
𐯉	𐯉	𐯉	𐯉, 𐯉
𐯊	𐯊	𐯊	𐯊
𐯋	𐯋	𐯋	𐯋, 𐯋
𐯌	𐯌	𐯌	𐯌
𐯍	𐯍	𐯍	𐯍, 𐯍
𐯎	𐯎	𐯎	𐯎
𐯏	𐯏	𐯏	𐯏, 𐯏
𐯐	𐯐	𐯐	𐯐
𐯑	𐯑	𐯑	𐯑, 𐯑
𐯒	𐯒	𐯒	𐯒
𐯓	𐯓	𐯓	𐯓, 𐯓
𐯔	𐯔	𐯔	𐯔
𐯕	𐯕	𐯕	𐯕, 𐯕
𐯖	𐯖	𐯖	𐯖
𐯗	𐯗	𐯗	𐯗, 𐯗
𐯘	𐯘	𐯘	𐯘
𐯙	𐯙	𐯙	𐯙, 𐯙
𐯚	𐯚	𐯚	𐯚
𐯛	𐯛	𐯛	𐯛, 𐯛
𐯜	𐯜	𐯜	𐯜
𐯝	𐯝	𐯝	𐯝, 𐯝
𐯞	𐯞	𐯞	𐯞
𐯟	𐯟	𐯟	𐯟, 𐯟
𐯠	𐯠	𐯠	𐯠
𐯡	𐯡	𐯡	𐯡, 𐯡
𐯢	𐯢	𐯢	𐯢
𐯣	𐯣	𐯣	𐯣, 𐯣
𐯤	𐯤	𐯤	𐯤
𐯥	𐯥	𐯥	𐯥, 𐯥
𐯦	𐯦	𐯦	𐯦
𐯧	𐯧	𐯧	𐯧, 𐯧
𐯨	𐯨	𐯨	𐯨
𐯩	𐯩	𐯩	𐯩, 𐯩
𐯪	𐯪	𐯪	𐯪
𐯫	𐯫	𐯫	𐯫, 𐯫
𐯬	𐯬	𐯬	𐯬
𐯭	𐯭	𐯭	𐯭, 𐯭
𐯮	𐯮	𐯮	𐯮
𐯯	𐯯	𐯯	𐯯, 𐯯
𐯰	𐯰	𐯰	𐯰
𐯱	𐯱	𐯱	𐯱, 𐯱
𐯲	𐯲	𐯲	𐯲
𐯳	𐯳	𐯳	𐯳, 𐯳
𐯴	𐯴	𐯴	𐯴
𐯵	𐯵	𐯵	𐯵, 𐯵
𐯶	𐯶	𐯶	𐯶
𐯷	𐯷	𐯷	𐯷, 𐯷
𐯸	𐯸	𐯸	𐯸
𐯹	𐯹	𐯹	𐯹, 𐯹
𐯺	𐯺	𐯺	𐯺
𐯻	𐯻	𐯻	𐯻, 𐯻
𐯼	𐯼	𐯼	𐯼
𐯽	𐯽	𐯽	𐯽, 𐯽
𐯾	𐯾	𐯾	𐯾
𐯿	𐯿	𐯿	𐯿, 𐯿
𐰀	𐰀	𐰀	𐰀
𐰁	𐰁	𐰁	𐰁, 𐰁
𐰂	𐰂	𐰂	𐰂
𐰃	𐰃	𐰃	𐰃, 𐰃
𐰄	𐰄	𐰄	𐰄
𐰅	𐰅	𐰅	𐰅, 𐰅
𐰆	𐰆	𐰆	𐰆
𐰇	𐰇	𐰇	𐰇, 𐰇
𐰈	𐰈	𐰈	𐰈
𐰉	𐰉	𐰉	𐰉, 𐰉
𐰊	𐰊	𐰊	𐰊
𐰋	𐰋	𐰋	𐰋, 𐰋
𐰌	𐰌	𐰌	𐰌
𐰍	𐰍	𐰍	𐰍, 𐰍
𐰎	𐰎	𐰎	𐰎
𐰏	𐰏	𐰏	𐰏, 𐰏
𐰐	𐰐	𐰐	𐰐
𐰑	𐰑	𐰑	𐰑, 𐰑
𐰒	𐰒	𐰒	𐰒
𐰓	𐰓	𐰓	𐰓, 𐰓
𐰔	𐰔	𐰔	𐰔
𐰕	𐰕	𐰕	𐰕, 𐰕
𐰖	𐰖	𐰖	𐰖
𐰗	𐰗	𐰗	𐰗, 𐰗
𐰘	𐰘	𐰘	𐰘
𐰙	𐰙	𐰙	𐰙, 𐰙
𐰚	𐰚	𐰚	𐰚
𐰛	𐰛	𐰛	𐰛, 𐰛
𐰜	𐰜	𐰜	𐰜
𐰝	𐰝	𐰝	𐰝, 𐰝
𐰞	𐰞	𐰞	𐰞
𐰟	𐰟	𐰟	𐰟, 𐰟
𐰠	𐰠	𐰠	𐰠
𐰡	𐰡	𐰡	𐰡, 𐰡
𐰢	𐰢	𐰢	𐰢
𐰣	𐰣	𐰣	𐰣, 𐰣
𐰤	𐰤	𐰤	𐰤
𐰥	𐰥	𐰥	𐰥, 𐰥
𐰦	𐰦	𐰦	𐰦
𐰧	𐰧	𐰧	𐰧, 𐰧
𐰨	𐰨	𐰨	𐰨
𐰩	𐰩	𐰩	𐰩, 𐰩
𐰪	𐰪	𐰪	𐰪
𐰫	𐰫	𐰫	𐰫, 𐰫
𐰬	𐰬	𐰬	𐰬
𐰭	𐰭	𐰭	𐰭, 𐰭
𐰮	𐰮	𐰮	𐰮
𐰯	𐰯	𐰯	𐰯, 𐰯
𐰰	𐰰	𐰰	𐰰
𐰱	𐰱	𐰱	𐰱, 𐰱
𐰲	𐰲	𐰲	𐰲
𐰳	𐰳	𐰳	𐰳, 𐰳
𐰴	𐰴	𐰴	𐰴
𐰵	𐰵	𐰵	𐰵, 𐰵
𐰶	𐰶	𐰶	𐰶
𐰷	𐰷	𐰷	𐰷, 𐰷
𐰸	𐰸	𐰸	𐰸
𐰹	𐰹	𐰹	𐰹, 𐰹
𐰺	𐰺	𐰺	𐰺
𐰻	𐰻	𐰻	𐰻, 𐰻
𐰼	𐰼	𐰼	𐰼
𐰽	𐰽	𐰽	𐰽, 𐰽
𐰾	𐰾	𐰾	𐰾
𐰿	𐰿	𐰿	𐰿, 𐰿
𐱀	𐱀	𐱀	𐱀
𐱁	𐱁	𐱁	𐱁, 𐱁
𐱂	𐱂	𐱂	𐱂
𐱃	𐱃	𐱃	𐱃, 𐱃
𐱄	𐱄	𐱄	𐱄
𐱅	𐱅	𐱅	𐱅, 𐱅
𐱆	𐱆	𐱆	𐱆
𐱇	𐱇	𐱇	𐱇, 𐱇
𐱈	𐱈	𐱈	𐱈
𐱉	𐱉	𐱉	𐱉, 𐱉
𐱊	𐱊	𐱊	𐱊
𐱋	𐱋	𐱋	𐱋, 𐱋
𐱌	𐱌	𐱌	𐱌
𐱍	𐱍	𐱍	𐱍, 𐱍
𐱎	𐱎	𐱎	𐱎
𐱏	𐱏	𐱏	𐱏, 𐱏
𐱐	𐱐	𐱐	𐱐
𐱑	𐱑	𐱑	𐱑, 𐱑
𐱒	𐱒	𐱒	𐱒
𐱓	𐱓	𐱓	𐱓, 𐱓
𐱔	𐱔	𐱔	𐱔
𐱕	𐱕	𐱕	𐱕, 𐱕
𐱖	𐱖	𐱖	𐱖
𐱗	𐱗	𐱗	𐱗, 𐱗
𐱘	𐱘	𐱘	𐱘
𐱙	𐱙	𐱙	𐱙, 𐱙
𐱚	𐱚	𐱚	𐱚
𐱛	𐱛	𐱛	

Buch-Pehlewi

Zeichenliste					
Nr.	Zeichen	Trans- position	Verbindungs- möglichkeit	Buchstabenwerte ¹⁾	Stilistische Varianten
1 bis 13 Zeichen, die zwischen Spatien vorkommen (Standardzeichen)					
Mit Neigung zur Einbeziehung von Ober- und Unterfeld (1 und 2):					
1	𐭡	M	beiderseits ²⁾	m, (<u>h</u>)	
2	𐭢	Z	beiderseits ³⁾	z, b, y	𐭣, 𐭤
Mit Neigung zur Einbeziehung des Unterfelds (3 bis 8):					
3	𐭥	<u>C</u>	—	c	𐭦, 𐭧
4	𐭨	<u>K</u>	von rechts	k	
5	𐭩	<u>W</u>	von rechts	wie Nr. 12	𐭪, mit M: 𐭫
6	𐭬	D	von rechts	d	
7	𐭭	B	von rechts	b, y, d	
8	𐭮	K	beiderseits	k	
Mit Oberlänge (9 bis 10):					
9	𐭯	L	beiderseits ³⁾	r, L ⁴⁾	𐭰
10	𐭱	<u>L</u>	beiderseits	R	𐭲, 𐭳
Im Mittelfeld (11 bis 13):					
11	𐭴	T	von rechts	t	
12	𐭵	W	von rechts	w, n, r, ' , ' ⁵⁾ , (<u>h</u>)	
13	𐭶	Y	beiderseits	y, d, g, b, z, (, h, s, s)	𐭷, 𐭸
14 bis 17 Zeichen im Mittelfeld, die der Verbindung nach einer Seite bedürfen (Hilfszeichen)					
14	𐭹	G	beiderseits	y, d, g, b, k, (s, s)	
15	𐭺	J	beiderseits	y, d, g, b, (, h, s)	
16	𐭻	I	beiderseits	y, d, g	
17	𐭼	C	von rechts	(c, p)	
18 bis 20 verschliffene Verbindungen					
18	𐭽	GYW _q ⁶⁾	—	'x ₁ ' ⁷⁾ ⁸⁾	
19	𐭾	YW _q ⁶⁾	nach links ⁹⁾	'x ₂ ' ⁷⁾	
20	𐭿	GḤ	von rechts	-yh	
¹⁾ Buchstaben, die aus mehr als einem Zeichen bestehen, in Klammern; über Ziffernwerte vgl. Bemerkungen am Schluß! ²⁾ 𐭡 auch von oben durchgezogen ³⁾ 𐭢 und 𐭣 auch nach unten durchgezogen ⁴⁾ 𐭰 = l, vgl. Buchstabenliste Nr. 11 mit Anm. 10! ⁵⁾ Schlußzeichen (Jod) ⁶⁾ q = Querstrich, d. i. Nr. 12: W ⁷⁾ Am Schluß von Verbal-Heterogrammen, mehrdeutig ⁸⁾ 𐭽 allein stehend = EYN ⁹⁾ Verbunden nur als -ynd in der Form 𐭽 YW _q I					

Figure 7a. Description of Book Pahlavi alphabet from a German source.

Buch-Pehlewi

Buchstaben in der Reihenfolge des aramäischen Alphabets					
Nr.	Name	Trans- literation	Realisierung	Zeichen	Diakritika, wenn richtig gesetzt
1	Aleph	'	Ⲁ	Y + Y	
			ⲁ	Y + J	
			Ⲃ	Y	
2	Beth	b	Ⲅ	B	
			ⲅ	Z	
			Ⲇ	J	
			ⲇ	G	
			Ⲉ	Y	
3	Gimel	g	Ⲋ	Y	mit ˘ = g
4	Daleth	d	ⲋ	I	mit ˘ = d
			Ⲍ	J	
			ⲍ	G	
			Ⲏ	D	mit ˘ = d
			ⲏ	B	
5	He	<u>H</u>	Ⲑ	MW	
			ⲑ	MMW	
6	Waw	w	Ⲓ	W	
			ⲓ	<u>W</u>	
7	Sajin	z	Ⲕ	Z	
			ⲕ	Y	
8	Cheth	h	wie Nr. 1 Aleph		mit ˘ = [ħ]
9	Jodh	y	ⲗ	Y	mit .. = y, [ȳ, ē] mit ˘ = [dʒ]
			Ⲙ	I	
			ⲙ	J	
			Ⲏ	G	
			ⲏ	Z	
			Ⲑ	B	
			ⲑ	W ⁹⁾	
10	Kaph	k	ⲓ	K	
			Ⲕ	<u>K</u>	
⁹⁾ Schlußzeichen					

Figure 7b. Description of Book Pahlavi alphabet from a German source.

Buch-Pehlewi

Nr.	Name	Trans- literation	Realisierung	Zeichen	Diakritika, wenn richtig gesetzt
11	Lamedh	r, L	ل	L	ل = l ¹⁰⁾
		$\bar{L}$	ل, ل	$\bar{L}$	
12	Mem	m	م	M	
13	Nun	n	wie Nr. 6 Waw		
14	Ssamech	s	س	G + G	mit ˙˙ = s
			س	G + Y	
15	ʿAjin	ʿ	wie Nr. 6 Waw		
16	Pe	p	پ	JC	
			پ	C	
17	Ssadhe	$c, \dot{S}$	س	$\bar{C}$	
			س	C	
			س	Y + W	
18	Resch	$\underline{r}$	wie Nr. 6 Waw		
19	Schin	$\dot{s}$	ش	JYY	mit ˙˙ = š
			ش	GY ¹¹⁾	
			ش	JY	
			ش	JYJ	
20	Taw	t	ت	T	mit ˙˙ = t
			ت	$\bar{T}$	
			ت	Y + W	

¹⁰⁾ ل in iranischen, ل in indischen Handschriften; l auch als ل, ل; vgl. Zeichenliste Nr. 9 mit Anm. 4!

¹¹⁾ Als š nur in indischen Handschriften

Bemerkungen über das Ziffernsystem:

1) Gleiche *Zeichengrundlage* wie bei Buchstaben; Zifferncharakter daher teilweise nur aus Kombination mehrerer Ziffern erkennbar oder aus dem Zusammenhang zu entnehmen.

2) *Ziffernfolge* absteigend nach Höhe der Zahlen, Rundzahlen zuerst, jedoch 20 höchster Ziffernwert unter 100; Einer verbunden in Gruppen von nicht mehr als vier: 11 = 10 + 1; 25 = 20 + 3 + 2; 37 = 20 + 10 + 4 + 3; 99 = 20 + 20 + 20 + 20 + 10 + 3 + 3 + 3

3) *Realisierung*: Einer isoliert oder final: ل B = 1, sonst nach links offener Haken im Mittelfeld halb- oder ganzstufig, im letzteren Fall mit Neigung zu Spitze auf der Unterlinie: ل YB, ل JB = 2, multiplizierend vor Hunderter oder Tausender auch mit Ansatzverdickung: ل ILZ = 1 × 100, ل سس = 6 × 1000; 10 = ل D, ل K, 1 × 20 = ل L; Zwanzigerverbindungen nach links offener Haken im Mittelfeld ganz-, aber auch halbstufig mit oder ohne Ansatzverdickung: ل GY = 20 + 20; ل GGY, ل YYY = 20 + 20 + 20; ل = 100

Figure 7c. Description of Book Pahlavi alphabet from a German source.

GRAMMAIRE PEHLEVIE.

CHAPITRE I.

SONS ET LETTRES.

I. La phonétique du pehlevi n'a point encore été systématisée ; trop de questions qui s'y rattachent, restent encore sans solution. Il semble cependant que l'on puisse admettre dans l'alphabet pehlevi les lettres et sons que voici :

VOYELLES : *a, i, u, e, o*, brèves et longues.

DIPHTHONGUES : *ai, au* ou *ae, ao*.

SEMI-VOYELLES : *r, l, y*.

CONSONNES. Gutturales : *k, g; kh, gh* (1), *h* (spirées).

Palatales : *c, j*.

Dentales : *t, d*.

Labiales : *p, f, b, v*.

Nasales : *n, m*.

Sibilantes : *z, s* (ç), *sh*.

II. L'alphabet pehlevi est d'origine sémitique. Il provient de l'araméen des derniers siècles de l'ère ancienne, par l'intermédiaire des alphabets sassanides. Il s'est formé et par

(1) *Gh* n'appartient pas réellement au pehlevi ; il représente 𐭪 avestique ou 𐭪.

Figure 8a. Description of Book Pahlavi from de Harlez 1880.

la régularisation des formes gravées et par une altération des caractères qui a ramené plusieurs d'entre eux à une même forme. Ainsi 𐭪 𐭪 et 𐭪 sont tous trois devenus 𐭪.

Les voyelles brèves médiales ne s'écrivent pas en général.

Les caractères pehlevi se tracent en allant de droite à gauche. L'alphabet pehlevi compte 14 lettres principales qui toutes, sauf la dernière, se retrouvent dans l'écriture avestique ; presque toutes représentent plusieurs sons.

Les voici avec leurs différentes valeurs, et dans l'ordre suivi par les lexicographes.

𐭪	𐭪	𐭪	𐭪	𐭪 ⁽¹⁾	𐭪	𐭪	𐭪 ⁽²⁾
a, â	b.	p, f.	t, d.	c, j, z.	r, l.	z.	s, ç.
h, kh.		v.	t.				
𐭪 (𐭪)	𐭪	𐭪	𐭪	𐭪	𐭪	𐭪	𐭪
sh, s.	gh.	k, g	m.	û, ô.		i, ê, y.	
				n, v, r, l.		g, d, j.	

On trouve encore 𐭪 pour *r* ou *l* final et 𐭪 pour *l*.

𐭪, 𐭪 et 𐭪 représentent souvent *a*, *u* et *i* brefs ; 𐭪 initial n'est parfois qu'une aspiration comme l'*élif* sémitique.

Les manuscrits de la seconde époque distinguent les différentes valeurs de 𐭪 par les signes diacritiques suivants : 𐭪 pour *i*, *ê*, *y* ; 𐭪 pour *j* ; 𐭪 pour *d*. La transcription nous dispensera de les employer.

Les ligatures formées par la jonction des lettres sont très nombreuses. La plupart représentent à la fois plusieurs groupes de lettres et grâce à la multiplicité de valeur des caractères, elles peuvent se lire de manières très différentes.

Dans un grand nombre de ces groupes les lettres ne sont que peu ou point déformées. Mais dans quelques uns, elles ne sont

(1) Parfois 𐭪. (2) Ou 𐭪 à la fin des mots et dans les groupes.

Figure 8b. Description of Book Pahlavi from de Harlez 1880.

presque point reconnaissables. Voici quelques principes qui pourront guider l'étudiant.

1° **u** initial d'un groupe perd le second trait et le premier resté seul se place en haut, à droite. Ex. **u** représente **u**. Le second trait s'allonge aussi en courbe. Ex. **uy** p. **uu**.

Final, **u** se rétrécit et se place au dessus, à gauche, comme dans **u**.

2° **d** initial n'est représenté que par un point terminant la courbe en haut à droite ou par un grossissement de la fin de cette courbe. Ainsi **u** représente **d** par le gros point final de droite et **u** par l'extrémité de gauche; la courbe inférieure unit les deux lettres.

3° La courbe inférieure que l'on rencontre dans beaucoup de groupes est tantôt une ligature, tantôt une partie intégrante d'une lettre. **uy** par exemple, peut représenter **ud** la courbe n'étant qu'une ligature, ou bien **du**, et alors la courbe fait partie de la deuxième lettre. (Voyez aussi 1°).

u initial ne garde souvent que le premier trait, le second s'allonge en courbe. Ainsi *çân* s'écrit **u** ou **u**.

4° **u** et **d** médiaux prennent la forme que l'on voit dans les petites courbes supérieures des groupes **u** *gân* ou *dân* et **u** *dîn*.

5° **u** final allonge le second trait en **u** et **d** final prend cette forme.

6° **y** reste distinct mais se joint à la lettre précédente et se courbe légèrement comme dans **uy**.

7° **u** remplace **u** et **u** final, joints à **d**.

Figure 8c. Description of Book Pahlavi from de Harlez 1880.

8° Certains lexicographes allongent inutilement le tableau des groupes pehlevis en y ajoutant les différentes valeurs qu'ils peuvent avoir selon les voyelles brèves qui entrent dans la formation des mots. C'est là une distinction inutile et dangereuse. Inutile, car on sait d'avance que **𐭮𐭥** *shn*, par ex. représente *shan*, *shen*, *shin*, *shun*, *shna*, *shne*, *shni*, etc.; dangereuse, parce qu'elle s'introduit dans les groupes un élément qui n'y est pas et pourrait donner à croire que ces voyelles y sont représentées d'une manière quelconque.

Voici les principaux groupes usités dans les manuscrits :

Tout groupe contenant une voyelle brève, médiale ou finale peut aussi se lire avec les autres voyelles brèves.

𐭮𐭥 pour **𐭮𐭥** se lit : *âh*, *ha*, *hâ*, *akh*, *âkh*, *khâ*, *hakh*.

— **𐭮𐭥** — *âî*.

— **𐭮𐭥** — *iyâ*, *îh*.

— **𐭮𐭥𐭥** — *ash*, *hash*, *âsh*.

— **𐭮𐭥𐭥** — *ish*, *êsh*, *gîh*.

𐭮𐭥 p. **𐭮𐭥** — *âh*, *khâ*, *hâ*, etc.

𐭮𐭥𐭥 p. **𐭮𐭥𐭥** — *âhâ*, *akhâ*, *khâha*, etc.

𐭮𐭥 p. **𐭮𐭥𐭥** — *îâv*.

— **𐭮𐭥𐭥** — *îâz*.

𐭮𐭥𐭥 p. **𐭮𐭥𐭥** — *âkht*, *khât*, *hât*, *hâd*, *hakht*.

𐭮𐭥𐭥 p. **𐭮𐭥𐭥** — *akht*, *âkht*, *hâta*, *hakht*.

— **𐭮𐭥𐭥𐭥** — *isht*, *îsht*, *yasht*, *dasht*.

— **𐭮𐭥𐭥𐭥** — *sât*, *sakht*, etc.

𐭮𐭥𐭥 p. **𐭮𐭥𐭥** — *âhm*, *âkham*, *khâm*, *hâm*.

— **𐭮𐭥𐭥𐭥** — *êshm*, *dashm*.

Figure 8d. Description of Book Pahlavi from de Harlez 1880.

𐭮	pour 𐭮𐭮 se lit :	<i>ahô, ahu, âhû, âhn, áán, akhv, ákhu, háva, hav, hán, khána, khân, etc.</i>
—	𐭮𐭮 —	<i>íán, êhó.</i>
—	𐭮𐭮𐭮 —	<i>ígán, êshva, îshn, dashn, dishn, gashn.</i>
—	𐭮𐭮𐭮 —	<i>sán, s-hû, sakha.</i>
𐭮	p. 𐭮 —	<i>ab, âb, hab.</i>
—	𐭮𐭮 —	<i>âi, hi khî.</i>
𐭮	p. 𐭮 —	<i>ap, âp, áf, haf, khef, av, áv.</i>
—	𐭮𐭮 —	<i>az, âz, haz, âj, âc.</i>
𐭮	p. 𐭮 —	<i>at, át, ad, ád, hat, khat, kht.</i>
—	𐭮𐭮 —	<i>dêt, gît.</i>
𐭮𐭮	p. 𐭮𐭮 —	<i>asb, hasb, etc.</i>
—	𐭮𐭮𐭮 —	<i>khayéb.</i>
—	𐭮𐭮𐭮 —	<i>aîy, khayé.</i>
𐭮	p. 𐭮 —	<i>am, áma, ham(a), hum, kham, etc.</i>
𐭮 ou 𐭮	p. 𐭮 —	1) <i>an, án, han, khan, khun.</i> 2) <i>aû, áu, áv, hu, hô, khû, khv.</i> 3) <i>hr, ár.</i> 4) <i>khal.</i>
𐭮	p. 𐭮 —	<i>aî, aé, ái, ay, hî, khî, ád, ag, áy, hag.</i>
—	𐭮𐭮 —	<i>és, das, gas, etc.</i>
𐭮	p. 𐭮 —	<i>aém, aîm, adm, agm, âgm, hîm, hêm, khîm, etc.</i>
𐭮	p. 𐭮 —	<i>aîn, áin, ayu, âîô, áên, hên, ayô, âgû, ágô, ádô, ádv, hên, hîn, khîn, khadû, khdû, khadv.</i>
—	𐭮𐭮 —	<i>diyô, édîn.</i>
—	𐭮𐭮 —	<i>îsva, yasn.</i>

Figure 8e. Description of Book Pahlavi from de Harlez 1880.

س	pour	س	se lit :	râ, rah, lâ.
س	p.	س	—	râm, rham, lâm, lakhm, etc.
	p.	س	—	raf, rap, rav, rp, rf, etc.
	—	س	—	rj, rc, rz, laj.
ر	p.	ر	—	rm, ram, rma, rim, lma, etc.
ر	p.	ر	—	raman, rmû.
ر	p.	ر	—	rû, rô, rv, rav; ran, runa, lû, lô, lav, lan, etc.
ر	p.	ر	—	rêm, rîm, radam, rgm, etc.
ز	p.	ز	—	zak, zek, etc.
س	p.	س	—	sb, sab.
	—	س	—	yêb, déb, igab, ggab.
	—	س	—	igy.
	—	س	—	deh, jeh, ikh, etc.
س	p.	س	—	spa, sip, suf.
	—	س	—	saz, sij.
	—	س	—	îz, êz, êj.
س	p.	س	—	st, sat, etc.
	—	س	—	yît, yêt, dît, gêt, etc.
س	p.	س	—	san, sen, sû, sa, sva, sav, sar.
	—	س	—	yîn, dîn, dêv, dên jîn, gîn, gdû, dgû, igû.
س	p.	س	—	sè, sî, sda, sag.
	—	س	—	is, das.
	—	س	—	dîg, etc.
س	p.	س	—	ih, iha, dâ, dah, dehe, dakh, gâ, já, jeh, yâ, yakh, etc.

Figure 8f. Description of Book Pahlavi from de Harlez 1880.

𐭮𐭮	pour	𐭮𐭮	se lit :	<i>shâ, shah, shkh, shaha.</i>
—	𐭮𐭮	—		<i>îhâ, yahâ, dahâ, gâh.</i>
—	𐭮𐭮	—		<i>gêha, yîh.</i>
—	𐭮𐭮	—		<i>yîsh.</i>
—	𐭮𐭮	—		<i>shîh.</i>
𐭮𐭮	p.	𐭮𐭮	—	<i>shân, shâu, shakhu.</i>
—	𐭮𐭮	—		<i>yadân.</i>
—	𐭮𐭮	—		<i>dehân, yahân.</i>
𐭮𐭮	p.	𐭮𐭮	—	<i>shab.</i>
—	𐭮𐭮	—		<i>yehab, yekhab, jâb, dahab, etc.</i>
𐭮𐭮	p.	𐭮𐭮	—	<i>sheb, shav.</i>
—	𐭮𐭮	—		<i>gâva, gâf, gâv.</i>
—	𐭮𐭮	—		<i>yâz, gâz, gâj, dâj.</i>
𐭮𐭮	p.	𐭮𐭮	—	<i>sht, shata.</i>
—	𐭮𐭮	—		<i>dât, dâte, gâta, jât, yât, ikhta, îhta, etc.</i>
𐭮𐭮	p.	𐭮𐭮	—	<i>shm, shama, etc.</i>
—	𐭮𐭮	—		<i>yâm, gâm, dâm. jâm, dakhm, etc.</i>
𐭮𐭮	p.	𐭮𐭮	—	<i>shn, shan, shl, shû, shô, shav, shva.</i>
—	𐭮𐭮	—		<i>yan, yân, îhû, îhô, yehu, yehen, gân, gâû, jân, jâva, dâna, dâva, dakhû, dahû, etc.</i>
𐭮𐭮	p.	𐭮𐭮	—	<i>shîn, shîr, shyô, shgô, shgan.</i>
—	𐭮𐭮	—		<i>dahyu, gâyo, êhîn.</i>
𐭮𐭮	p.	𐭮𐭮	—	<i>shâh, shaha.</i>
—	𐭮𐭮	—		<i>shedâ.</i>
—	𐭮𐭮	—		<i>yâhâ.</i>

Figure 8g. Description of Book Pahlavi from de Harlez 1880.

𐭠𐭡	p. 𐭠𐭡	—	maman, mman, mmû, etc.
𐭠	p. 𐭠	—	man, mû, mar, mil? etc.
𐭠𐭡𐭢	p. 𐭠𐭡 ou 𐭠𐭡𐭢	se lit :	madam ou mehim.
𐭠𐭡	p. 𐭠𐭡	—	ib, êb, db, dab, gab. yî, gî.
𐭠	p. 𐭠	—	iv, îf.
	— 𐭠𐭡	—	ic, îc, îj, dej, gaj, if, îz, daz, gaz, yaz, etc.
𐭠𐭡	p. 𐭠𐭡	—	îta, êt, id, yet, dat, gat, dta, gta, etc.
𐭠𐭡	p. 𐭠𐭡 ou 𐭠𐭡𐭢	se lit :	îk, yak, îg, dak, jig.
𐭠𐭡	p. 𐭠𐭡	—	îm, êm, gm, dm, dam, yem.
𐭠	p. 𐭠	—	ên, in, în, iv, îva, etc. îr, etc. dô, du, dn, dan, den, dav, dar, gan, gô, gû, jan, jô, jar, jar, etc. yô, yû, etc.
𐭠𐭡	p. 𐭠𐭡	—	êdm, gadm, jadm, yadm, yêm.
	— 𐭠𐭡	—	sma, sam.
𐭠	p. 𐭠𐭡	—	îga, dî, daga, gî, etc.
𐭠𐭡		—	yên, yîn.

Souvent 𐭠 se sépare en deux parties. La seconde prend la forme de la courbe initiale et s'unit au groupe suivant :

Ex. 𐭠𐭡 pour 𐭠𐭡; 𐭠𐭡𐭢 pour 𐭠𐭡𐭢; 𐭠𐭡𐭢𐭣 pour 𐭠𐭡𐭢𐭣.

Figure 8h. Description of Book Pahlavi from de Harlez 1880.

Row 10B: BOOK PAHLAVI

	10B4	10B5	10B6
0	𐭢𐭣	𐭢	𐭠𐭡𐭢
1	𐭢𐭣𐭤	𐭢	𐭠𐭡𐭢𐭣
2	𐭢	𐭢	
3	𐭢𐭣𐭤		
4	𐭢		
5	𐭢𐭣𐭤		𐭢
6	𐭢		𐭢𐭣𐭤
7	𐭢𐭣𐭤		𐭢𐭣𐭤
8	𐭢𐭣𐭤		𐭢𐭣𐭤
9	𐭢𐭣𐭤		𐭢𐭣𐭤
A	𐭢𐭣𐭤	𐭢𐭣𐭤	𐭢𐭣𐭤
B	𐭢𐭣𐭤	𐭢𐭣𐭤	𐭢𐭣𐭤
C	𐭢𐭣𐭤	𐭢𐭣𐭤	𐭢𐭣𐭤
D	𐭢𐭣𐭤	𐭢𐭣𐭤	𐭢𐭣𐭤
E	𐭢𐭣𐭤	𐭢𐭣𐭤	𐭢𐭣𐭤
F	𐭢𐭣𐭤	𐭢𐭣𐭤	𐭢𐭣𐭤

hex	Name
40	BOOK PAHLAVI LETTER ALEPH-HET
41	BOOK PAHLAVI LETTER BETH
42	BOOK PAHLAVI LETTER GIMEL-DALETH-YODH
43	BOOK PAHLAVI LETTER HE
44	BOOK PAHLAVI LETTER WAW-NUN-AYIN-RESH
45	BOOK PAHLAVI LETTER ZAYIN
46	BOOK PAHLAVI LETTER KAPH
47	BOOK PAHLAVI LETTER GHAPH
48	BOOK PAHLAVI LETTER LAMEDH
49	BOOK PAHLAVI LETTER LHAMEDH
4A	BOOK PAHLAVI LETTER MEM-QOPH
4B	BOOK PAHLAVI LETTER SAMEKH
4C	BOOK PAHLAVI LETTER PE
4D	BOOK PAHLAVI LETTER SADHE
4E	BOOK PAHLAVI LETTER SHIN
4F	BOOK PAHLAVI LETTER TAW
50	BOOK PAHLAVI ARCHIGRAPHHEME EAR
51	BOOK PAHLAVI ARCHIGRAPHHEME ELBOW
52	BOOK PAHLAVI ARCHIGRAPHHEME BELLY
53	(This position shall not be used)
54	(This position shall not be used)
55	(This position shall not be used)
56	(This position shall not be used)
57	(This position shall not be used)
58	(This position shall not be used)
59	(This position shall not be used)
5A	BOOK PAHLAVI COMBINING GIMEL
5B	BOOK PAHLAVI COMBINING DALETH
5C	BOOK PAHLAVI COMBINING YODH
5D	BOOK PAHLAVI COMBINING SAMEKH
5E	BOOK PAHLAVI COMBINING SHIN
5F	BOOK PAHLAVI KASHIDA
60	BOOK PAHLAVI ABBREVIATION TAA
61	BOOK PAHLAVI LOGOGRAM TURNED AHREMAN
62	(This position shall not be used)
63	(This position shall not be used)
64	(This position shall not be used)
65	BOOK PAHLAVI NUMBER ONE
66	BOOK PAHLAVI NUMBER TWO
67	BOOK PAHLAVI NUMBER THREE
68	BOOK PAHLAVI NUMBER FOUR
69	BOOK PAHLAVI NUMBER TEN
6A	BOOK PAHLAVI NUMBER TWENTY
6B	BOOK PAHLAVI NUMBER FORTY
6C	BOOK PAHLAVI NUMBER SIXTY
6D	BOOK PAHLAVI NUMBER EIGHTY
6E	BOOK PAHLAVI NUMBER ONE HUNDRED
6F	BOOK PAHLAVI NUMBER ONE THOUSAND

A. Administrative

1. Title

Proposal to encode the Avestan script in the BMP of the UCS

2. Requester's name

Michael Everson and Roozbeh Pournader

3. Requester type (Member body/Liaison/Individual contribution)

Individual contribution.

4. Submission date

2007-07-30

5. Requester's reference (if applicable)

6. Choose one of the following:

6a. This is a complete proposal

Yes.

6b. More information will be provided later

No.

B. Technical – General

1. Choose one of the following:

1a. This proposal is for a new script (set of characters)

Yes.

1b. Proposed name of script

Avestan.

1c. The proposal is for addition of character(s) to an existing block

No.

1d. Name of the existing block

2. Number of characters in proposal

64.

3. Proposed category (A-Contemporary; B.1-Specialized (small collection); B.2-Specialized (large collection); C-Major extinct; D-Attested extinct; E-Minor extinct; F-Archaic Hieroglyphic or Ideographic; G-Obscure or questionable usage symbols)

Category C.

4a. Is a repertoire including character names provided?

Yes.

4b. If YES, are the names in accordance with the “character naming guidelines” in Annex L of P&P document?

Yes.

4c. Are the character shapes attached in a legible form suitable for review?

Yes.

5a. Who will provide the appropriate computerized font (ordered preference: True Type, or PostScript format) for publishing the standard?

Michael Everson.

5b. If available now, identify source(s) for the font (include address, e-mail, ftp-site, etc.) and indicate the tools used:

Michael Everson, Fontographer.

6a. Are references (to other character sets, dictionaries, descriptive texts etc.) provided?

Yes.

6b. Are published examples of use (such as samples from newspapers, magazines, or other sources) of proposed characters attached?

Yes.

7. Does the proposal address other aspects of character data processing (if applicable) such as input, presentation, sorting, searching, indexing, transliteration etc. (if yes please enclose information)?

Yes.

8. Submitters are invited to provide any additional information about Properties of the proposed Character(s) or Script that will assist in correct understanding of and correct linguistic processing of the proposed character(s) or script. Examples of such properties are: Casing information, Numeric information, Currency information, Display behaviour information such as line breaks, widths etc., Combining behaviour, Spacing behaviour, Directional behaviour, Default Collation behaviour, relevance in Mark Up contexts, Compatibility equivalence and other Unicode normalization related information. See the Unicode standard at <http://www.unicode.org> for such information on other scripts. Also see Unicode Character Database <http://www.unicode.org/Public/UNIDATA/UnicodeCharacterDatabase.html> and associated Unicode Technical Reports for information needed for consideration by the Unicode Technical Committee for inclusion in the Unicode Standard.

See above.

C. Technical – Justification

1. Has this proposal for addition of character(s) been submitted before? If YES, explain.

Yes. See N2556, N1684.

2a. Has contact been made to members of the user community (for example: National Body, user groups of the script or characters, other experts, etc.)?

Yes.

2b. If YES, with whom?

Hassan Rezai Baghbidi, Hossein Masoumi Hamedani (Iranian Academy of Sciences), Jost Gippert, Desmond Durkin-Meisterernst, Günter Schweiger

2c. If YES, available relevant documents

<http://titus.fkidg1.uni-frankfurt.de/unicode/iranian/3tagung.htm>

3. Information on the user community for the proposed characters (for example: size, demographics, information technology use, or publishing use) is included?

Zoroastrians, Iranianists and other scholars.

4a. The context of use for the proposed characters (type of use; common or rare)

Used liturgically and by scholars.

4b. Reference

5a. Are the proposed characters in current use by the user community?

Yes.

5b. If YES, where?

Religious and scholarly publications.

6a. After giving due considerations to the principles in the P&P document must the proposed characters be entirely in the BMP?

Yes.

6b. If YES, is a rationale provided?

Yes.

6c. If YES, reference

Accordance with the Roadmap. Avestan is used in modern Zoroastrian religion.

7. Should the proposed characters be kept together in a contiguous range (rather than being scattered)?

No.

8a. Can any of the proposed characters be considered a presentation form of an existing character or character sequence?

No.

8b. If YES, is a rationale for its inclusion provided?

8c. If YES, reference

9a. Can any of the proposed characters be encoded using a composed character sequence of either existing characters or other proposed characters?

No.

9b. If YES, is a rationale for its inclusion provided?

9c. If YES, reference

10a. Can any of the proposed character(s) be considered to be similar (in appearance or function) to an existing character?

No.

10b. If YES, is a rationale for its inclusion provided?

10c. If YES, reference

11a. Does the proposal include use of combining characters and/or use of composite sequences (see clauses 4.12 and 4.14 in ISO/IEC 10646-1: 2000)?

No.

11b. If YES, is a rationale for such use provided?

11c. If YES, reference

11d. Is a list of composite sequences and their corresponding glyph images (graphic symbols) provided?

No.

11e. If YES, reference

12a. Does the proposal contain characters with any special properties such as control function or similar semantics?

No.

12b. If YES, describe in detail (include attachment if necessary)

13a. Does the proposal contain any Ideographic compatibility character(s)?

No.

13b. If YES, is the equivalent corresponding unified ideographic character(s) identified?