

Accredited Standards Committee¹

NCITS, Information Processing Systems

Doc: **L2/97-103**
Date: 21 May 1997
Author: John H. Jenkins
Apple Computer, Inc.
2 Infinite Loop
M/S 302-2IS
Cupertino CA 95014
Phone: (408) 974-6276
Fax: (408) 862-4566
e-mail: jenkins@apple.com

Proposal to add Shavian to Unicode/ISO-IEC 10646

Background

The playwright George Bernard Shaw (1856-1950) was an outspoken critic of the idiosyncrasies of English orthography. In his will, he directed that Britain's Public Trustee seek out and publish an alphabet of no fewer than forty letters to provide for the phonetic spelling of English. The alphabet finally selected was designed by Kingsley Read and is variously known as Shavian, Shaw's alphabet, and the Proposed British Alphabet.

Also in accordance with Shaw's will, an edition of his play, *Androcles and the Lion* was published and distributed to libraries, containing the text in both the standard Latin alphabet and in Shavian.

As with other attempts at spelling reform in English, the alphabet has met with little success. It has, nonetheless, its advocates and users. There are at least two implementations for computers, for example, and a standard implementation of Unicode's user zone includes an encoding of Shavian.

The normative version of Shavian is taken to be the version in *Androcles and the Lion*.

Structure

The alphabet consists of forty-eight letters and one punctuation mark. The letters have no case. The digits and other punctuation marks are the same as Latin. The one additional punctuation mark is a "name mark," used to indicate proper nouns. The letter names are intended to be indicative of their sounds; thus the sound /p/ is represented by the letter PEEP (ᵶ).

The first forty letters are divided into four groups of ten. The first ten and second ten are 180-degree rotations of one another; letters of the third and fourth often show a similar relationship of shape.

¹ Operating under the procedures of the American National Standards Institute

NCITS Secretariat, Information Technology Industry Council

1250 Eye Street NW, Suite 200, Washington DC 20005-3922 Telephone: 202-737-8888 (press 1 twice) Fax: 202-638-4922

The first ten letters are "tall" letters (which ascend above the x-height) and generally represent unvoiced consonants. The next ten letters are "deep" letters (which descend below the baseline) and generally represent voiced consonants. The next twenty are the vowels and liquids. Again, each of these letters usually has a close phonetic relationship to the letter in its matching set of ten.

The remaining eight letters are technically ligatures, the first six involving vowels plus /r/. Since ligation isn't optional, these eight letters are included in the proposal.

Proposal

This proposal is based on the version of the Shaw alphabet found in the ConScript Unicode user-zone registry. It allocates three columns for the forty-eight letters and uses U+00B7 MIDDLE DOT for the name mark.

The proposed location for the encoding of the Shaw alphabet is on Plane 1, off the BMP, and in the same block as the Deseret Alphabet. Shavian's limited use makes this a reasonable location. This is also the position indicated for Shavian in the current Road Map proposal. We indicate in **boldface** sound in each letter's name which the letter represents. This is a convention adopted from *Androcles and the Lion* but which need not be maintained in Unicode and ISO/IEC 10646.

0001 xx00	ʅ	SHAVIAN LETTER PEEP
0001 xx01	1	SHAVIAN LETTER TOT
0001 xx02	ɔ	SHAVIAN LETTER KICK
0001 xx03	J	SHAVIAN LETTER FEE
0001 xx04	ð	SHAVIAN LETTER THIGH
0001 xx05	ʂ	SHAVIAN LETTER SO
0001 xx06	ʟ	SHAVIAN LETTER SURE
0001 xx07	ʑ	SHAVIAN LETTER CHURCH
0001 xx08	\	SHAVIAN LETTER YEA
0001 xx09	ʈ	SHAVIAN LETTER HUNG
0001 xx0A	ʟ	SHAVIAN LETTER BIB
0001 xx0B	ɹ	SHAVIAN LETTER DEAD
0001 xx0C	ɹ	SHAVIAN LETTER GAG
0001 xx0D	ɹ	SHAVIAN LETTER VOW
0001 xx0E	ɹ	SHAVIAN LETTER THEY
0001 xx0F	ɹ	SHAVIAN LETTER ZOO
0001 xx10	ɹ	SHAVIAN LETTER MEASURE
0001 xx11	ɹ	SHAVIAN LETTER JUDGE
0001 xx12	/	SHAVIAN LETTER WOE
0001 xx13	ɹ	SHAVIAN LETTER HA-HA
0001 xx14	c	SHAVIAN LETTER LOLL
0001 xx15	ɹ	SHAVIAN LETTER MIME
0001 xx16	ɹ	SHAVIAN LETTER IF

0001 xx17	ℓ	SHAVIAN LETTER EGG
0001 xx18	ℓ	SHAVIAN LETTER ASH
0001 xx19	ℓ	SHAVIAN LETTER ADO
0001 xx1A	ℓ	SHAVIAN LETTER ON
0001 xx1B	ℓ	SHAVIAN LETTER WOOL
0001 xx1C	ℓ	SHAVIAN LETTER OUT
0001 xx1D	ℓ	SHAVIAN LETTER AH
0001 xx1E	ℓ	SHAVIAN LETTER ROAR
0001 xx1F	ℓ	SHAVIAN LETTER RUN
0001 xx20	ℓ	SHAVIAN LETTER EAT
0001 xx21	ℓ	SHAVIAN LETTER AGE
0001 xx22	ℓ	SHAVIAN LETTER ICE
0001 xx23	ℓ	SHAVIAN LETTER UP
0001 xx24	ℓ	SHAVIAN LETTER OAK
0001 xx25	ℓ	SHAVIAN LETTER OOZE
0001 xx26	ℓ	SHAVIAN LETTER OIL
0001 xx27	ℓ	SHAVIAN LETTER AWE
0001 xx28	ℓ	SHAVIAN LETTER ARE
0001 xx29	ℓ	SHAVIAN LETTER OR
0001 xx2A	ℓ	SHAVIAN LETTER AIR
0001 xx2B	ℓ	SHAVIAN LETTER ERR
0001 xx2C	ℓ	SHAVIAN LETTER ARRAY
0001 xx2D	ℓ	SHAVIAN LETTER EAR
0001 xx2E	ℓ	SHAVIAN LETTER IAN
0001 xx2F	ℓ	SHAVIAN LETTER YEW

A table of the forty-eight proposed letters follows:

	0001 xx0	0001 xx1	0001 xx2
0	ℓ	ℓ	ℓ
1	ℓ	ℓ	ℓ
2	ℓ	ℓ	ℓ
3	ℓ	ℓ	ℓ
4	ℓ	ℓ	ℓ
5	ℓ	ℓ	ℓ
6	ℓ	ℓ	ℓ

7	Ʒ	Ƶ	ƶ
8	ƶ	ƶ	ƶ
9	ƶ	ƶ	ƶ
A	ƶ	ƶ	ƶ
B	ƶ	ƶ	ƶ
C	ƶ	ƶ	ƶ
D	ƶ	ƶ	ƶ
E	ƶ	ƶ	ƶ
F	ƶ	ƶ	ƶ

Issues

The problem of collation is not addressed by the alphabet's designers.

References

Cowan, John, and Everson, Michael. 1997. Shavian ConScript Unicode Standard. <http://www.indigo.ie/egt/standards/csur/shavian.html>.

Crystal, David, ed. 1987. *The Cambridge Encyclopedia of Language*. Cambridge: University Press.

DeMeyere, Ross. 1997. About Shavian. <http://www.demeyere.com/Shavian/info.html>.

Shaw, George Bernard. 1962. *Androcles and the Lion; an Old Fable Renovated by Bernard Shaw, with a Parallel Text in Shaw's Alphabet to be Read in Conjunction Showing its Economies in Writing and Reading*. Harmondsworth: Penguin Books

Examples

A key to the alphabet from *Androcles and the Lion*:

The Shaw Alphabet Reading Key

The letters are classified as Tall, Deep, Short, and Compound.
Beneath each letter is its full name: its *sound* is shown in **bold** type.

Tall :										
	peep	tot	kick	fee	thigh	so	sure	church	yea	hung
Deep :										
	bib	dead	gag	vow	they	zoo	measure	judge	woe	ha-ha
Short :										
	loll	mime	if	egg	ash	ado	on	wool	out	ah
	roar	nun	eat	age	ice	up	oak	ooze	oil	awe
Compound :										
	are	or	air	err	array	ear	Ian	yew		

The four most frequent words are represented by single letters: the *q*, of *f*, and **, to *l*.

Proper names may be distinguished by a preceding 'Namer' dot: e.g. *·Dof*, Rome.

Punctuation and numerals are unchanged. Learn the alphabet *in pairs*, as listed for Writers overleaf.

התאריך: 20.08.2018

קרסנו

[ספרות: חשבון שווק, שווק פ' שווק, חשבון
הן שווק.

[illegible][illegible][illegible]

The announcement for the Tenth International Unicode Conference in Shavian (note the lack of the naming mark in the text—*mea culpa, mea culpa, mea maxima culpa*):

לכך, שג"כ + ר"ח:
 קודם קודם / ר"ח

[illegible]

ԱՐ ԲՈՐ ԱՆՏ 1 ԴԵ, 11 ՏԻՎՏ
ԱՆԻՈՐ

የአገሪቱን የፍትሕ ፍጥነት ማረጋገጥና የፍትሕ ፍጥነት ማረጋገጥ.

המחברת מודה לרבותן על סיוען, תמיכתן, עזרתן וסובלנותן.

האגודה, וכל האגודות, עוסקות בשיפור חיינו, וכל האגודות, וכל האגודות, עוסקות בשיפור חיינו.

[illegible][illegible]