ISO/IEC JTC1/SC2/WG2 N4262
L2/12-138
2012-04-29

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

Doc Type: **Working Group Document**
Title: **Proposal to encode "Unifon" and other characters in the UCS**
Source: **UC Berkeley Script Encoding Initiative (Universal Scripts Project)**
Author **Michael Everson**
Status: **Individual Contribution**
Date: **2012-04-29**

1. Introduction to Unifon. Unifon was developed as an auxiliary "phonetic" alphabet designed to facilitate access to literacy to English-speaking children, by presenting to them a writing system that worked by sound. Tests showed that children were able to learn to read rather quickly using this system, and, having made that breakthrough, were able to transition to traditional English orthography relatively easily. Unifon was developed in the 1950s by Dr John R. Malone, an economist and newspaper equipment consultant who became interested in phonetic writing while consulting with the Bendix Corporation, which was interested in questions of aviation communication. That work was abandoned when the International Air Transport Association selected English as the language of international airline communications in 1957. But Malone's interest in phonetic writing resurfaced when his young son complained about difficulties learning to read. From about 1960 to the 1980s, Margaret S. Ratz used Unifon to teach first-graders at Principia College in Elsah, Illinois. A variety of teaching materials exist using Unifon. From the 1974 to his death in 1993 John M. Culkin, a specialist in media studies, also promoted Unifon.

Of greater significance is the use made of Unifon in the 1970s and 1980s to write Native American languages. Unifon was adapted principally by Tom Parsons of Humboldt State University to provide a practical orthography for several the Hupa, Yurok, Tolowa, and Karok languages. These orthographies were used for a nunber of years and although other orthographies are used for these languages now, many valuable documents using Unifon exist which should be able to take advantage of UCS encoding.

2. Structure. Unifon is a bicameral script written left to right. Most Unifon text is written in ALL CAPITAL LETTERS, but the system as developed and described does permit the use of casing pairs; when casing is used, the lower-case forms are conventionally (that is, *always*) written in SMALL CAPITAL LETTERS. Unifon uses 40 characters when used for writing English; a number of additional characters were used for the Native American languages, and a few characters were used in earlier versions of Unifon but were later replaced by other characters.

3. Encoding model. Because of the considerable overlap between many Unifon letters and the Latin script, Unifon should be treated as a set of extensions to the Latin script. A large number of Unifon characters should be unified with existing Latin characters. As noted above, when Unifon is used as a casing script, it is always intended to be displayed as styled text, THAT IS, IN SMALL-CAPS. This has some implications for the design of lower-case letters for the code charts, but that should not be particularly problematic if standard design principles are applied.

4. Combining diacritical marks. Generic diacritical marks are used in Unifon orthography. In Tolowa, U+0301 COMBINING ACUTE ACCENT is used to indicate stress, and both U+0304 COMBINING MACRON and U+0331 COMBINING MACRON BELOW are used to "harden" and "soften" the sound made by X [x].

5. Disunificiations. A number of characters used in practical orthographies and in phonetic transcriptions are related to some of the characters used in Unifon, which is why these are included in this proposal. Several of these are related to letters encoded in the Greek script. Evidence for these is given before the further presentation of Unifon characters.

Ᵽ	A7AE	LATIN CAPITAL LETTER BETA
ᵽ	A7AF	LATIN SMALL LETTER BETA
		• used in Gabonese orthographies
Ɵ	A7B0	LATIN CAPITAL LETTER THETA
ɵ	A7B1	LATIN SMALL LETTER THETA
		• used in Unifon and Tuscarora orthographies
Χ	A7B2	LATIN CAPITAL LETTER CHI
χ	A7B3	LATIN SMALL LETTER CHI
		• used in Lepsius phonetic orthography
Ꞷ	A7B4	LATIN CAPITAL LETTER OMEGA
ꞷ	A7B5	LATIN SMALL LETTER OMEGA
		• used in Gabonese orthographies
ᴉ	A7B6	LATIN CAPITAL LETTER SMALL CAPITAL I
		x 026A ɪ LATIN LETTER SMALL CAPITAL I
		• used in Unifon and Gabonese orthographies

5.1 LATIN LETTER BETA was arguably disunified from Greek by the devisers of the IPA. David Abercrombie describes this in his 1967 *Elements of General Phonetics*:

> A good source from which letters can be borrowed is the Greek alphabet, and β γ ɛ θ ɸ χ, for example, have been made use of for centuries in roman-based phonetic notations. Borrowed Greek letters are sometimes redesigned so as to fit in with the general appearance of roman letters. The preceding six characters, for example, have for this reason been modified as follows: β ɣ ɛ θ ɸ χ. The Cyrillic alphabet can also offer possible new characters, such as ь ч, and the script form *ш*.

Here we see the already-disunified ɣ LATIN LETTER GAMMA, ɛ LATIN LETTER OPEN E, and ɸ LATIN LETTER PHI, which are quite distinct from the usual Greek γ, ε, and φ (shown here also in Times). The normal shape of β GREEK LETTER BETA has no serif on its lower descender, and as can be seen

here, the shape of the Latinized IPA β BETA is based more on the ß SHARP S than the original Greek letter. In Daniel Jones' 1932 *An outline of English phonetics*, an even less "Greek-like" beta can be seen (Jones was Assistant Secretary of the International Phonetic Association from 1907 to 1927, Secretary from 1927 to 1949, and President from 1950 to 1967):

> **806.** The sound **w** causes difficulty to many foreigners, especially to Germans. They generally replace it by a different kind of bi-labial fricative, namely one in which the lips are kept flat instead of being rounded and pushed forward, and in which the tongue is in a neutral position instead of being raised at the back. The phonetic symbol for this consonant is **ß**. Its lip-position is the same as that of **Φ** (Fig. 89). It is a sound intermediate in acoustic effect between **w** and **v**; it is very frequently heard in German words like *Quelle* ˈ**kßelə** or ˈ**kvelə**, *zwei* **tßai** or **tsvai**. Sometimes foreigners replace **w** by **v**.

The LATIN BETA in lower-case and upper-case form has been found in *Revue Gabonaises des Sciences de l'Homme*, No. 2, 1990, p.113. The usage is based on the "Alphabet scientifique des langues du Gabon" (ASG) first published 1989, which was followed by the "Orthographe des langues du Gabon" intended for the educational system in 1999. The unique Latin capital form Ɓ is unknown for Greek B:

ʃ	ʃ	ísʃékè	"hérisson"	*yinzɛbi*
t	T	tètètè	"tuméfié"	*omyɛnɛ-mpoŋgwɛ*
u	U	tsúlù	"source"	*lekaniŋi*
ụ	Ụ	ékù̱rì	"piège"	*faŋ-ntumu*
v	V	úvèɣə	"donner"	*yisangu*
β	Ɓ	íβéβè	"voler"	*bɛŋga*
w	W	wàɣêndà	"étrangers"	*yeßia*
ẉ	Ẉ	njẉé	"chef"	*faŋ-ntumu*

From the same journal, page 193, The word *ɣeßoße* in all caps (*ƔEƁOƁE*):

Ɣ E Ɓ O Ɓ E

The unique Latin capital form meets one of the major criteria for disunification. Whether the existence of this character (or several of the others proposed below) would affect the recommendations of the International Phonetic Association is a matter for the eventual decision of the Association. Certainly much IPA text currently uses the Greek β. Much probably uses the Latin ß, and certainly much IPA text also currently uses Latin B for the same character in pre-UCS encoded fonts, which are still unfortunately more widespread than one might wish. Support for a disunification has been given by John Wells, currently a member of the Council of the Association:

> Michael Everson correctly identifies a number of reasons to advocate the disunification of the Latin letters beta, theta, and chi from their Greek versions. If this happened, as IPA symbols we would use the Latin versions rather than the Greek ones.
>
> He quotes briefly, without identifying the source, from the IPA 1949 *Principles* booklet. Here, more fully, is what is says there (*The Principles of the International Phonetic Association*, pages 1-2). Although unattributed, these are clearly Daniel Jones's words.

> (*c*) The non-roman letters of the International Phonetic Alphabet have been designed as far as possible to harmonise well with the roman letters. The Association does not recognise makeshift letters ; it recognises only letters which have been carefully cut so as to be in harmony with the other letters. For instance, the Greek letters included in the International Alphabet are cut in roman adaptations. Thus, since the ordinary shape of the Greek letter *β* does not harmonise with roman type, in the International Phonetic Alphabet it is given the form **β**.

Note the very clear intention to treat IPA **θ** (vertical) as distinct from Greek theta (typically oblique). Greek letters are to be incorporated into the IPA only as roman [sic] adaptations.

> And of the two form of Greek theta, *θ* and *ϑ*, it has been necessary to choose the first (in vertical form), since the second cannot be made to harmonise with roman letters.

As Jones says, Greek theta has an alternative form, *ϑ*. This is encoded at U+03D1, whereas ordinary *θ* is at U+03B8.

In English printed texts that mix the Latin and Greek scripts, the Greek letters are typically oblique, the Latin ones upright. The purpose is to distinguish clearly between the two scripts (whereas the IPA wants everything in the same script). Here is an example, from Abbott and Mansfield's *Primer of Greek Grammar* (my copy printed in 1949).

EXAMPLES.

ἠχώ (ο), *ἠ*, *echo*, like *πειθώ* ; *δμώς* (ω), *ὁ*, *slave*, like *ἥρως*.

51. *Accentuation in Declension.*

(1.) The accent remains, so far as possible, on the syllable which is accentuated in the nom. case. Thus *γένος*, gen. *γένους*, *χελιδών*, gen. *χελιδόνος*.

(2.) The genitive and dative of monosyllabic nominatives are generally accented on the last syllable in all numbers, *e.g.* *θήρ*, gen. *θηρός*, *θηρί*, *θηροῖν*, *θηρῶν*, *θηρσί*. Short syllables are oxytone, long are perispomenon. So also *γυνή*, *woman*. *παίδων* from *παῖς*, and *ὦτων* from *οὖς*, are exceptions.

I think disunification of Latin and Greek beta, theta, chi would be a good thing.

If it was the intention of the founders of the IPA to borrow Latin letters distinctly into the IPA from Greek, it would seem that disunification of some but not all of them is a long-standing mistake on the part of the early designers of the UCS.

5.2 LATIN LETTER THETA has been discussed above by John Wells, citing Daniel Jones. Its unification with Greek is problematic for Unifon and for other orthographies because of the extremely peculiar casing relationships which are evidently acceptable for Greek, but not acceptable for Latin:

0398 Θ GREEK CAPITAL LETTER THETA lower-cases to 03B8 θ GREEK SMALL LETTER THETA
03B8 θ GREEK SMALL LETTER THETA upper-cases & title-cases to 0398 Θ GREEK CAPITAL LETTER THETA
03D1 ϑ GREEK THETA SYMBOL upper-cases and title-cases to 0398 Θ GREEK CAPITAL LETTER THETA
03F4 ϴ GREEK CAPITAL THETA SYMBOL lower-cases to 03B8 θ GREEK SMALL LETTER THETA

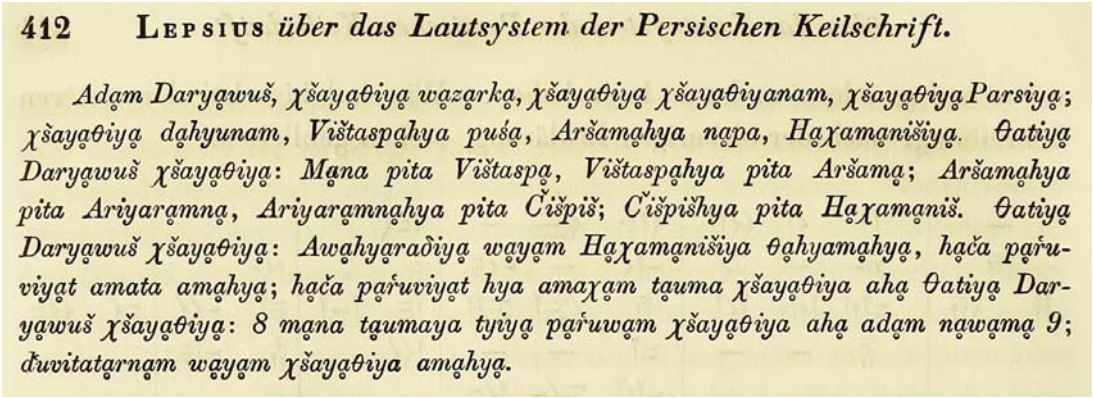
So essentially, 0398 Θ has two lower-case letters, 03B8 θ and 03D1 ϑ, and 03B8 θ has two upper-case letters, 0398 Θ and 03F4 ϴ. But LATIN THETA cases ϴ with θ, which isn't a reliable, *or reversible* pairing in Greek. Here is an example of this from *The Tuscarora Language, Beginner Program*:

ϴaʔneθwe:kih
Wakaʔneθwe:kih
Ruʔneθwe:kih
Yakuʔneθwe:kih
Yuʔneθwe:kih

Taking the four letters Θ θ ϑ ϴ and applying (in Quark XPress) “all caps” to them yields Θ Θ Θ Θ, and applying “small caps” to then yields Θ ʘ ʘ Θ.

In any event, the correct casing relation for Tuscarora, for early Unifon, and for the early Phonotypic alphabet is Θ θ (Θ Θ, Θ ʘ), not Θ θ (Θ Θ, Θ ʘ). The behaviour of the Greek letters is incorrect for these Latin orthographies. Again, this argues for a disunification.

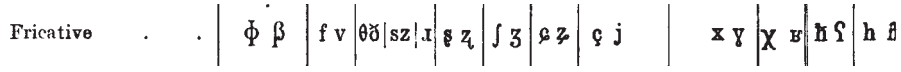
The important 19th-century linguist Richard Lepsius also designed his LATIN THETA quite differently from the usual way in which GREEK THETAS are drawn; they are like Θ θ but with the horizontal line extending somewhat past either side, thus ϴ θ:



412 LEPSIUS *über das Lautsystem der Persischen Keilschrift.*

Adąm Daryąwuš, χšayąθiyą wązarką, χšayąθiyą χšayąθiyanam, χšayąθiyąParsiyą; χšayąθiyą dąhyunam, Vištaspąhya puśą, Aršamąhya nąpa, Haχamąnišiyą. θatiyą Daryąwuš χšayąθiyą: Mąna pita Vištaspą, Vištaspąhya pita Aršamą; Aršamąhya pita Ariyaramną, Ariyaramnąhya pita Čišpiš; Čišpišhya pita Haχamąniš. θatiyą Daryąwuš χšayąθiyą: Awąhyaraδiyą wąyam Haχamąnišiya θąhyamąhyą, hąča pąṙuviyąt amata amąhyą; hąča pąṙuviyąt hya amaχąm tąuma χšayąθiya ahą θatiyą Daryąwuš χšayąθiyą: 8 mąna tąumaya tyiyą pąṙuwąm χšayąθiya ahą adąm nąwąmą 9; δuvitatąrnąm wąyąm χšayąθiya amąhyą.

5.3 LATIN LETTER CHI was arguably disunified from Greek by the devisers of the IPA. This can be seen in the 1949 chart:



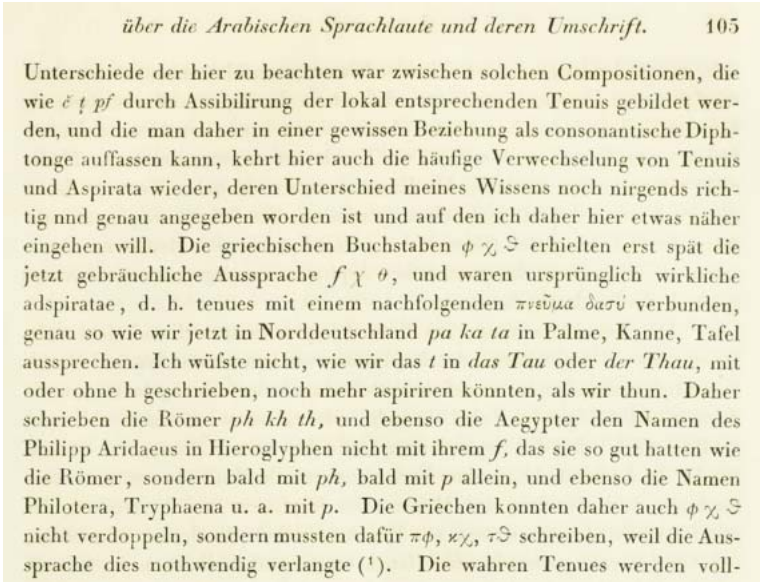
Here the Latin serifs and the thick leg of the LATIN CHI go from top right to bottom left, and the curvy leg goes from top left to bottom right. Here is a comparison between LATIN X, LATIN STRETCHED X (used in Teuthonista notation), GREEK CHI, and LATIN CHI:

X x χ ~ X χ ~ X χ

The weight of LATIN SMALL LETTER X and STRETCHED X is on the \ leg, but the weight of GREEK SMALL LETTER CHI and LATIN SMALL LETTER CHI (at least as used in the IPA) is on the / leg. The size of LATIN CAPITAL LETTER CHI is also unique and unknown in Greek. This is similar to CYRILLIC CAPITAL KU which can have the shapes Ԛ or Q, while the former is not an acceptable variant for LATIN CAPITAL Q. Here is an example of Lepsius distinguishing GREEK CAPITAL CHI and LATIN CAPITAL CHI:

wird das *m* von *χnum*,
ın Χνοῦβι ϑεῷ geweiht; ir

Lepsius also distinguishes the three Latin letters *f χ θ* from the Greek letters *ϕ χ θ*:



über die Arabischen Sprachlaute und deren Umschrift. 105

Unterschiede der hier zu beachten war zwischen solchen Compositionen, die wie *ĕ ţ pf* durch Assibilirung der lokal entsprechenden Tenuis gebildet werden, und die man daher in einer gewissen Beziehung als consonantische Diphtonge auffassen kann, kehrt hier auch die häufige Verwechselung von Tenuis und Aspirata wieder, deren Unterschied meines Wissens noch nirgends richtig nnd genau angegeben worden ist und auf den ich daher hier etwas näher eingehen will. Die griechischen Buchstaben *ϕ χ ϑ* erhielten erst spät die jetzt gebräuchliche Aussprache *f χ θ*, und waren ursprünglich wirkliche adspiratae, d. h. tenues mit einem nachfolgenden *πνεῦμα δασύ* verbunden, genau so wie wir jetzt in Norddeutschland *pa ka ta* in Palme, Kanne, Tafel aussprechen. Ich wüfste nicht, wie wir das *t* in *das Tau* oder *der Thau*, mit oder ohne h geschrieben, noch mehr aspiriren könnten, als wir thun. Daher schrieben die Römer *ph kh th*, und ebenso die Aegypter den Namen des Philipp Aridaeus in Hieroglyphen nicht mit ihrem *f*, das sie so gut hatten wie die Römer, sondern bald mit *ph*, bald mit *p* allein, und ebenso die Namen Philotera, Tryphaena u. a. mit *p*. Die Griechen konnten daher auch *ϕ χ ϑ* nicht verdoppeln, sondern mussten dafür *πϕ, κχ, τϑ* schreiben, weil die Aussprache dies nothwendig verlangte (¹). Die wahren Tenues werden voll-

5.4 LATIN LETTER OMEGA is used in several Kulango language publications from Bondounkou or its area, including an alphabet primer ("syllabaire"), four liturgy books (a missal), a funeral book, and two catechism books. All were published in the 1990s, most probably by the diocese of Bondounkou. This example comes from *Les funérailles chrétiennes*, p17:

(Y) Bɪ tɛsɛ Jezu, yɷgɔ bɛrɛ pɪɪ pɛɛ, waa kprɛ-bɛ lɛ bɔ nyī
ūguogo vɛɛ. Dɛ lɛ tɪɪ ɷ sɷmsɛ ... hɷ̃ɷ daa gbigo hɔɔ.
ɷ kyɛrɛ: mɷm mī̱ɪ lɛ mɷ̃, kpɷ̃kpɛ'rɛ hɷ̃ hɛ saakɔ hɔɔ-rɪ, zɛɛ-ɛ
lɛ kyɛ-lɛ, lɛ hɷ̃ nyī ī̱yakɔ ɷ kyī̱ŋɔ ariginā-nɪ. Prɛɷ̃ ɷɷ yi lɛ
bɷŋɷnɪ drunyā, lɛ ɷ kprɛ-gɛ. Bɪɪ dalɪ-ɷ lɛ ɷ kpele bɪ yɷɷkɔ-rɪ
ɷ sɪra-rɔ ....

It can be seen that Latin ɷ ɷ are not the same as Greek Ω ω. The Times glyph for ɷ was based on the IPA character ɷ LATIN SMALL LETTER CLOSED OMEGA.

5.5 LATIN LETTER SMALL CAPITAL I is also used in several Kulango language publications from Bondounkou. This example (from Psalm 118) comes from *Ī̱ dɛ bī̱ dalı bıtɛsɛ*, p 10:

Alleluya!
Ī̱ kprɛ Yego bɔ gyasole, gyɪgalɛɪ hɷ̃ kyɛrɛ,
gyɪgalɛɪ bɔ korigyo tuben haa ti-ɪ.
Izraɛl bɔ yɔgɔnɪyɷgɔ, ī̱ dɷ-kɛ:
bɔ korigyo tuben haa ti-ɪ.

6. Alphabets using Unifon characters. A number of Unifon letters should be unified with existing Latin characters. These have capital and small letter forms, and these are given below. Again, when

Unifon is printed using casing pairs, the lower case is conventionally (that is, *always*) presented in small caps, so in the presentation here both upper and lower case are given, and also the lower case shown in small caps. These are given in a typical Unifon font style.

The 40-letter alphabet presently used for English is as follows:

A 40-letter alphabet previously used for English was as follows:

A 40-letter alphabet previously used for English was as follows:

The 40-letter Shaw-Malone Forty-Phoneme Alphabet for English was as follows:

The 33-letter alphabet used for Hupa was as follows:

The 26-letter alphabet used for Karuk was as follows:

The 30-letter alphabet used for Tolowa was as follows:

A 32-letter alphabet used for Yurok was as follows:

A 42-letter "Indian Unifon Single-Sound Alphabet" is given as follows:

A number of other letters (not listed in the alphabets above) derive from earlier versions of Unifon:

7.1 Unifon letters unified with existing characters. Of the 66 letters used in the various Unifon alphabets, 43 of them—about two-thirds—can be unified with existing letters. Note that *none* of the small-cap letters are encoded modifier letters: they are small-caps styled forms of ordinary small letters.

Letter name	Capital	small	SMALL-CAP	UNIFON
LATIN LETTER A	A	a	ᴀ	[æ]
LATIN LETTER TURNED V	Λ	ʌ	ʌ	[ɒ]
LATIN LETTER B	B	b	ʙ	[b]
LATIN LETTER C	C	c	ᴄ	[s]
LATIN LETTER C WITH STROKE	Ȼ	ȼ	ȼ	[tʃ]
LATIN LETTER REVERSED C	Ɔ	ɔ	ɔ	[tʃ]
LATIN LETTER D	D	d	ᴅ	[d]
LATIN LETTER E	E	e	ᴇ	[ɛ]
LATIN LETTER TURNED E	Ǝ	ə	ǝ	[ə]
LATIN LETTER F	F	f	ꜰ	[f]
LATIN LETTER G	G	g	ɢ	[g]
LATIN LETTER H	H	h	ʜ	[h]
LATIN LETTER J	J	j	ᴊ	[dʒ]
LATIN LETTER K	K	k	ᴋ	[k]
LATIN LETTER L	L	l	ʟ	[l]
LATIN LETTER M	M	m	ᴍ	[m]
LATIN LETTER N	N	n	ɴ	[n]
LATIN LETTER O	O	o	ᴏ	[ɔ]
LATIN LETTER O WITH STROKE	Ø	ø	ø	[ʊ]
LATIN LETTER P	P	p	ᴘ	[p]
LATIN LETTER R	R	r	ʀ	[r]
LATIN LETTER S	S	s	ꜱ	[s]
LATIN LETTER T	T	t	ᴛ	[t]
LATIN LETTER U	U	u	ᴜ	[ʌ]
LATIN LETTER V	V	v	ᴠ	[v]
LATIN LETTER W	W	w	ᴡ	[w]
LATIN LETTER X	X	x	x	[x]
LATIN LETTER Y	Y	y	ʏ	[j]
LATIN LETTER Z WITH STROKE	Ƶ	ƶ	ƶ	[ʒ]

7.2. New characters for Unifon. A number of Unifon letters should be added in order to support Unifon in its various stages of development. The character names here are tentative, thought it was thought better to try to use descriptive names rather than the Shavian-like "ICE, EYES, BIT, RING" for instance.

Letter name	Capital	small	SMALL-CAP	
LATIN LETTER CLOSED TURNED V	Ʌ	ʌ	ʌ	[eɪ]
LATIN LETTER REVERSED-E E	Ɛ	ɶ	ɛ	[iː]
LATIN LETTER SCHWA WITH HOOK	ɚ	[ɚ]	ɚ	[ɚ]
LATIN LETTER SMALL CAPITAL I	Ɪ	ɪ	ɪ	[ɪ]
LATIN LETTER I WITH STROKE AND BASELINE	Ɨ	ɨ	ɨ	[aɪ]
LATIN LETTER I WITH SERIFED STROKE	Ɨ	ɨ	ɨ	[eɪ]
LATIN LETTER OVERTURNED WINEGLASS	Ƃ	ƃ	ƃ	[eɪ]
LATIN LETTER REVERSED N WITH BENT RIGHT LEG	И	и	и	[ŋ]
LATIN LETTER O WITH BASELINE	Ω	ω	ω	[oʊ]
LATIN LETTER O WITH VERTICAL BAR	Ɵ	ɵ	ɵ	[ʊ]
LATIN LETTER O WITH LOW VERTICAL BAR	ɷ	ɷ	ɷ	[aʊ]

LATIN LETTER O WITH HIGH VERTICAL BAR	Ɵ	ɵ	ɵ	[ɔɪ]
LATIN LETTER OY	Ƣ	ƣ	ƣ	[ɔɪ]
LATIN LETTER S WITH STROKE	ẞ	ß	ß	[ʃ]
LATIN LETTER THE	Ƕ	ƕ	ƕ	[θ]
LATIN LETTER THETA	Θ	θ	ϴ	[θ]
LATIN LETTER DHE	Ƭ	ƭ	ƭ	[ð]
LATIN LETTER TURNED T	Ʇ	[ʇ]	ʇ	[ð]
LATIN LETTER U WITH BASELINE	Ʉ	ʉ	ʉ	[uː]
LATIN LETTER CLOSED U	Ʊ	ʊ	ʊ	[ju]
LATIN LETTER U WITH VERTICAL STROKE	Ʒ	ʒ	ʒ	[uː]
LATIN LETTER REVERSED Z	Ƹ	ƹ	ƹ	[z]
LATIN LETTER CHE	Ч	ч	ч	[tʃ]
LATIN LETTER TLE	Ƕ	ƕ	ƕ	[eɪ]

7.3.1 SCHWA WITH HOOK. A few issues are worth pointing out. One option of encoding ƎR is as LATIN LETTER SCHWA WITH HOOK, being the capital letter of the already-encoded ɚ. But strictly speaking, since Ǝə TURNED E and Əə SCHWA are distinct, one might expect a case pair Ə˞ə˞ for SCHWA WITH HOOK. In that case two new characters would be needed, either TURNED-E R Ǝ with either ɐɾ or ɚ or ɐʀ or ɐʁ or ʁ as the lower-case. Here is a larger version of these characters for easier discussion:

Ǝə ~ Ə˞ə˞ ~ ƎRɐɾɚɐʀɐʁƎʁ

7.3.2 REVERSED-E E. It was assumed that the ligature was of *ɘ* and *e*, rather than of *ə* and *e*, since the intended sound is *ee* [iː]:

ɘe ≠ əe

9. Bibliography.

Anderson, Kenneth C. 2007. *The Case for a Sensible Alphabet*. http://www.unifon.org/documents/The%20Case%20for%20a%20Logical%20Alphabet.pdf

Anderson, Kenneth C. 2007. *Learn Unifon—Spell the Sounds!*. http://www.unifon.org/documents/The%20Spelling%20Revolution.pdf

Bennett, Ruth. 1981. *Xo-xul-woth: He is dug up*. Arcata. Humboldt State University.

Bennett, Ruth and Carolyn Risling Shaw. 1984. *Basketmaking among the Karuk*. Arcata. Humboldt State University.

Bennett, Ruth. 1984. *Hupa Consonants*. Arcata. Humboldt State University.

Bennett, Ruth. 1984. *Hupa Nature Coloring Book*. Arcata. Humboldt State University.

Bennett, Ruth. 1985. *Hupa Natural Resources Dictionary*. Arcata. Humboldt State University.

Bennett, Ruth. 1985. *New Hupa Spelling Book*. Arcata. Humboldt State University.

Bennett, Ruth et al. 1986. *Hupa Numbers*. Arcata. Humboldt State University.

Bennett, Ruth et al. 1986. *Karuk Numbers*. Arcata. Humboldt State University.

Bennett, Ruth et al. 1986. *Yurok Sentences*. Arcata. Humboldt State University.

Bennett, Ruth et al. 1986. *Yurok Vocabulary*. Arcata. Humboldt State University.

Bennett, Ruth. 1987. *Hupa Activity Book*. Arcata. Humboldt State University.

Bennett, Ruth. 1987. *Cooperative Learning with a Computer in a Native Languaage Class*. Arcata. Humboldt State University.

Bennett, Ruth. 1987. *Teaching and Learning with Computers: A Method for American Indian Bilingual Classrooms*. Arcata. Humboldt State University.

Bommelyn, Loren and B. Humphrey. 1983. *The Tolowa Language*. Arcata. Humboldt State University.

Carroll, Lewis. [2012, in press]. *Alice's Adventures in Wonderland in the Unifon Alphabet: ALISIZ ADVENCƏRZ IN WUNDƏRLAND*. Cathair na Mart: Evertype. ISBN 978-1-78201-003-6.

Davis, Shan. 1986. *Karuk Lesson Book*. Edited by Ruth Bennett. Arcata. Humboldt State University.

Guy, Gladys, and Shan Davis. 1986. *Karuk Lesson Book*. Arcata. Humboldt State University.

Malone, John R. 1960. "Do we need a new alphabet?" in *Chicago Sunday Sun-Times*, 1960-05-29.

Rudes, Blair A. 1999. *Tuscarora-English/English-Tuscarora dictionary*. University of Toronto Press, ISBN 0-8020-4336-4

9. 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) in respect of the Unifon encoding. Any views, findings, conclusions or recommendations expressed in this publication do not necessarily reflect those of the National Endowment for the Humanities.

CAP.	small	SM.CAP.	CAP.	small	SM.CAP.	Name
A	a	A	A	a	A	LATIN LETTER A
Δ	ɐ	Δ	Δ	ɐ	Δ	LATIN LETTER CLOSED TURNED V
Λ	ʌ	Λ	Λ	ʌ	Λ	LATIN LETTER TURNED V
B	b	B	B	b	B	LATIN LETTER B
C	c	C	C	c	C	LATIN LETTER C
Ȼ	ȼ	Ȼ	Ȼ	ȼ	Ȼ	LATIN LETTER C WITH STROKE
Ɔ	ɔ	Ɔ	Ɔ	ɔ	Ɔ	LATIN LETTER REVERSED C
D	d	D	D	d	D	LATIN LETTER D
E	e	E	E	e	E	LATIN LETTER E
Ɛ	ɶ	Ɛ	Ɛ	ɶ	Ɛ	LATIN LETTER REVERSED-E E
Ǝ	ə	Ǝ	Ǝ	ə	Ǝ	LATIN LETTER AFRICAN E
ɚ	ɚ	ɚ	ɚ	ɚ	ɚ	LATIN LETTER SCHWA WITH HOOK
F	f	F	F	f	F	LATIN LETTER F
G	g	G	G	g	G	LATIN LETTER G
H	h	H	H	h	H	LATIN LETTER H
Ɪ	ɪ	Ɪ	Ɪ	ɪ	Ɪ	LATIN LETTER SMALL CAPITAL I
Ɨ	ɨ	Ɨ	Ɨ	ɨ	Ɨ	LATIN LETTER I WITH STROKE AND BASELINE
Ɨ	ɨ	Ɨ	Ɨ	ɨ	Ɨ	LATIN LETTER I WITH SERIFED STROKE
⅄	⅄	⅄	⅄	⅄	⅄	LATIN LETTER OVERTURNED WINEGLASS
J	j	J	J	j	J	LATIN LETTER J
K	k	K	K	k	K	LATIN LETTER K
L	l	L	L	l	L	LATIN LETTER L
Ł	ł	Ł	Ł	ł	Ł	LATIN LETTER L WITH HOR. BAR
M	m	M	M	m	M	LATIN LETTER M
N	n	N	N	n	N	LATIN LETTER N
Ɲ	ɲ	Ɲ	Ɲ	ɲ	Ɲ	LATIN LETTER REVERSED N WITH BENT RIGHT LEG
O	o	O	O	o	O	LATIN LETTER O
Ɑ	ɑ	Ɑ	Ɑ	ɑ	Ɑ	LATIN LETTER O WITH BASELINE

CAP.	small	SM.CAP.	CAP.	small	SM.CAP.	Name
Ꝋ	ꝋ	ᴏ	Ꝋ	ꝋ	ᴏ	LATIN LETTER O WITH VERTICAL BAR
Ꝍ	ꝍ	ᴏ	Ꝍ	ꝍ	ᴏ	LATIN LETTER O WITH LOW VERTICAL BAR
Ꝏ	ꝏ	ᴏ	Ꝏ	ꝏ	ᴏ	LATIN LETTER O WITH HIGH VERTICAL BAR
Ꝍꞵ	Ꝍꞵ	Ꝍꞵ	Ꝍꞵ	Ꝍꞵ	Ꝍꞵ	LATIN LETTER OY
Ø	ø	Ø	Ø	ø	Ø	LATIN LETTER O WITH STROKE
P	p	P	P	p	P	LATIN LETTER P
R	r	R	R	r	R	LATIN LETTER R
S	s	S	S	s	S	LATIN LETTER S
ꞩ	ꞩ	ꞩ	ꞩ	ꞩ	ꞩ	LATIN LETTER S WITH STROKE
T	t	T	T	t	T	LATIN LETTER T
Ꞥ	ꞥ	Ꞥ	Ꞥ	ꞥ	Ꞥ	LATIN LETTER DHE
Ꝋ	ꝋ	ᴏ	Ꝋ	ꝋ	ᴏ	LATIN LETTER THETA
Ꞧ	ꞧ	Ꞧ	Ꞧ	ꞧ	Ꞧ	LATIN LETTER THE
Ʇ	Ʇ	Ʇ	Ʇ	Ʇ	Ʇ	LATIN LETTER TURNED T
U	u	U	U	u	U	LATIN LETTER U
Ʉ	ʉ	Ʉ	Ʉ	ʉ	Ʉ	LATIN LETTER U WITH BASELINE
ᵾ	ᵾ	ᵾ	ᵾ	ᵾ	ᵾ	LATIN LETTER CLOSED U
Ꝟ	ꝟ	Ꝟ	Ꝟ	ꝟ	Ꝟ	LATIN LETTER U WITH VERTICAL BAR
V	v	V	V	v	V	LATIN LETTER V
W	w	W	W	w	W	LATIN LETTER W
X	x	X	X	x	X	LATIN LETTER X
Y	y	Y	Y	y	Y	LATIN LETTER Y
Ƶ	ƶ	Ƶ	Ƶ	ƶ	Ƶ	LATIN LETTER Z WTH STROKE
Ꞓ	Ꞓ	Ꞓ	Ꞓ	Ꞓ	Ꞓ	LATIN LETTER REVERSED Z
Ꞗ	ꞗ	Ꞗ	Ꞗ	ꞗ	Ꞗ	LATIN LETTER CHE
Ꞅ	ꞅ	Ꞅ	Ꞅ	ꞅ	Ꞅ	LATIN LETTER TLE

Figures.

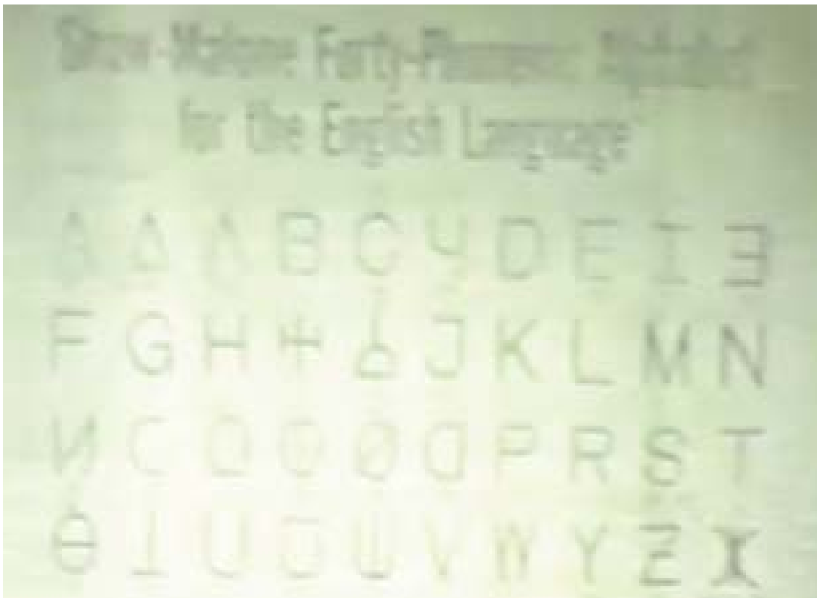


Figure 1. Example of the first published version, the Shaw-Malone Forty-Phoneme Alphabet, taken as a screen shot of a video in which John Malone was giving a lecture about the alphabet. The alphabet given is: AΔΛBCЧDEIƎFGHǂᐃJKLMNИOΩΦØΘPRSTΘ⊥UŪШVWYƵX

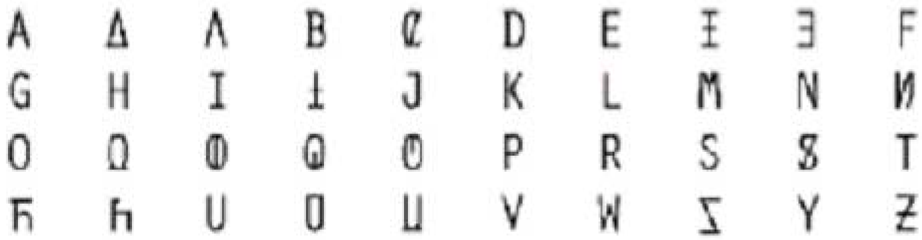


Figure 2. Example of an intermediate version of Unifon. The alphabet given is: AΔΛBȻDEꞮƎFGHI±JKLMNИOΩΦΘΦPRS$TĦĦUŪꓴVWΣYƵ

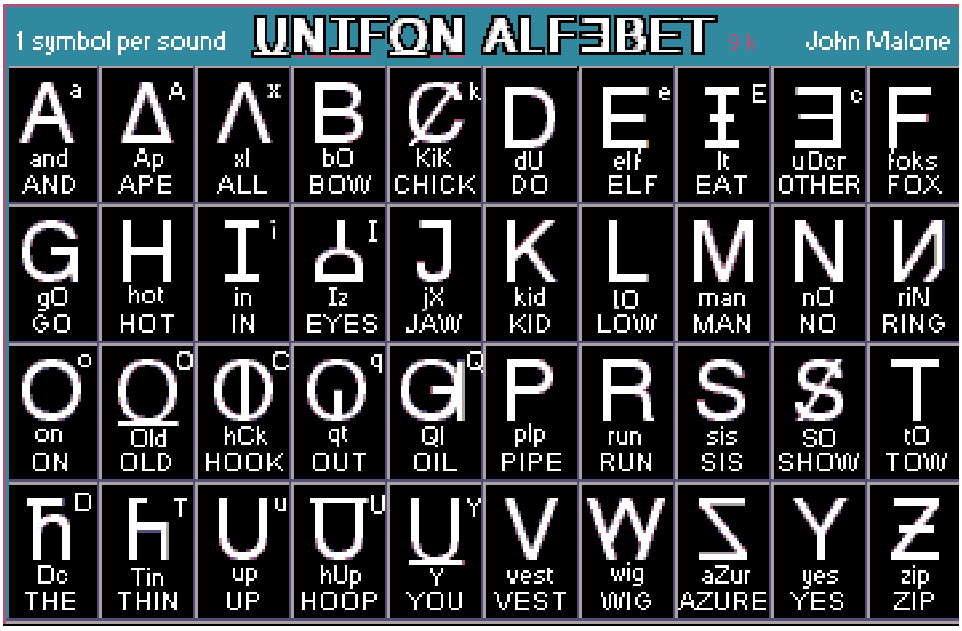


Figure 3. Example of an intermediate version of Unifon. The alphabet given is: AΔΛBȻDEꞮƎFGHIᐃJKLMNИOΩΦΘGPRS$TĦĦUŪꓴVWΣYƵ

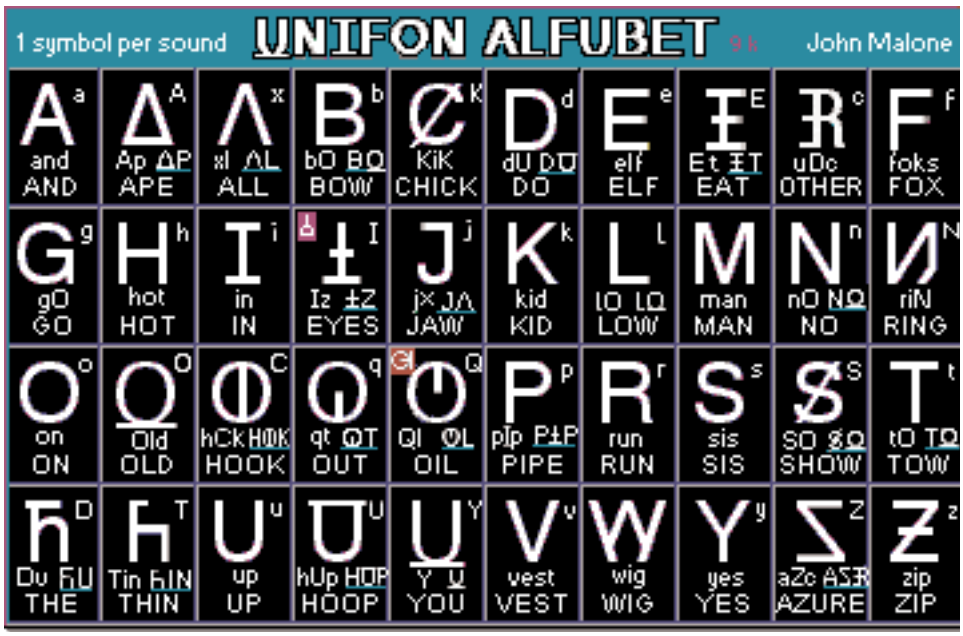


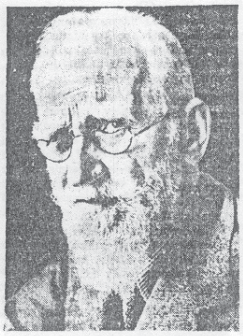
Figure 4. Example of the final version of Unifon. The alphabet given is:
AΔΛBȻDEƎƎFGHI±JKLMNИOQΘΘΟPRSƧTĦHUŪUVWƩYƵ

Figure 5. Example of the final version of Unifon. The alphabet given is:
AΔΛBȻDEƎƎFGHI±JKLMNИOQΘΘΟPRSƧTĦHUŪUVWƩYƵ

EDITORIALS
COLUMNS
CHICAGO SUNDAY
SUN-TIMES
May 29, 1960
NEWS REVIEW
SPECIAL ARTICLES
Section Two

MY FAIR LANGUAGE

Do We Need A New Alphabet?



GEORGE BERNARD SHAW
A legacy for reform of the language

Poor Eliza Doolittle! The winsome Cockney flower girl of "My Fair Lady," waged a desperate and delightful struggle to master the King's English for Prof. 'Enry 'Iggins. Anne Rogers played role in Chicago.

By John R. Malone

Two and a half years ago, I spent one of the pleasantest evenings of my life listening to and looking at the delights of "My Fair Lady" . . . written over George Bernard Shaw's "Pygmalion." As I'm sure you all know, the story turns on the efforts of a phonetics professor, Henry Higgins, to change the language structure of a little Cockney flower girl to that of "upper class" Albertian English—and the cultural transformation which this brings about for her.

Early in the show Prof. 'Enry 'Iggins sings his plaintive lament: "Oh why can't the English learn to speak?" while his precious guttersnipe, Eliza Doolittle, distorts the noble tongue beyond recognition. I believe he knows why.

Being of Irish extraction and having something of a sympathy for the Shavian disdain of doing things the same way as others do because that is the custom of the past, the complaint of Prof. Higgins set me to thinking . . . right out there in the theater in plain sight of the beautiful stage sets.

AND WHILE THE beautiful, lilting music-hall tunes filled the evening, my mind went back to the 6-year-old at home who was having all sorts of trouble with spelling English words. It was neither consistent nor logical as his numbers were. Sometimes he could "sound out" a word; but mostly spelling was a grab bag. Getting the proper letters for a word was pure chance.

"Why can't the English . . . ?" Suddenly I saw clearly the answer to the professor's rhetorical question. Obviously they couldn't learn to speak because they hadn't learned to spell. And they hadn't learned to spell because they *couldn't*. They didn't have their own adequate alphabet and there weren't enough letters in the Latin alphabet with which they had been burdened. Eureka!

As the English are a hybrid people made up of Celts, Druids, Romans, Jutes, Saxons, Angles, Normans, Danes and Norsemen, their language is made up of elements of all these with remnants of Gaelic, Flemish and Platt-deutsch (a low German tongue) and spelled with the leftovers of an alphabet left by Julius Caesar before the time of Christ. This alphabet was later reimposed upon them by way of the Latin-based Christian church.

SINCE EARLY HEBREW and Phoenician times, the written language has generally been related rather closely to the sounds of the language. And the Phoenician and classic Greek languages were excellent examples of this. However, these languages used basically 20 to 25 or so sounds or phonemes, and their alphabets or language symbols had to be within this range. The Latin adaptation made a few ground rules for sounding vowels in one of two ways, and using some letters such as I and V for both vowels and consonants; but Latin, too, was relatively simple, using 22 letters to represent from 26 to 28 sounds.

But English! As a problem in linguistics it is plenty tough. First, it is made up of from 39 to 44 sounds coming from some of the sources indicated above, many of which are not Latin or Greek at all. Using the already inadequate Latin alphabet of 22 letters to represent these sounds made it even tougher. To do this at all satisfactorily at least four letters (J, U, W and Y) have been added. And all sorts of consistent and inconsistent ground rules have been made for giving different letters and combinations of letters different sounds in different words. This was done hundreds of years ago, with or without good cause, and today we are stuck with the whole kaboodle of them.

IN THE OLDEN DAYS every scribe or clerk had his own feelings about spelling because there were few dictionaries and no printers. A big stew pot of inconsistent rules grew up to cover the sonorous, expressive collection of words from all Europe and elsewhere, which became known as "English." Only a people so patient and stubborn as the English would have even tried to make a pattern of spelling out of such a mixed-up situation. But "muddling through" solved it . . . in a way.

Then the typesetters were brought to English shores by William Caxton, the first English printer. From Holland and Flanders he brought them and their type fonts to put "Reynard the Fox" into print — the first English typeset work. At that time, they used a Latin alphabet with 24 letters and had no fixed rules for setting up English. So these Dutch and Flemish printers made their own rules as they went or tried to use the continental rules if

Continued on Next Page

G. B. SHAW'S WILL —THE BACKGROUND

Infuriated by such manifest absurdities as pronouncing ph to sound like f, George Bernard Shaw, the great Irish playwright and critic, applied his trenchant wit to our alphabetical aggravations and came up with a characteristically iconoclastic solution: Invent a written language with enough characters so that each letter would designate a specific invariable sound. To this end he willed that much of his estate go toward an award for a more adequate, economical and forthright orthography of the English language. The probate court allowed $23,240 for prizes to contestants and the expense of cutting type and setting up an edition of Shaw's "Androcles and the Lion" in a doubly typeset volume—half new alphabet and half the orthodox English. This, in the proviso of Shaw's will, is to be distributed to 13,000 public libraries in English-speaking lands.

Among the ten finalists in this competition was one of the 60 Americans who were among the 467 entrants from all over the world. He is a Chicago advertising executive whose background includes wide experience as newspaperman, magazine contributor, consulting economist, Army captain with overseas service and assistant professor at his alma mater, the University of Kansas. He is working on his doctorate at the University of Chicago. He is John R. Malone, 46, whose many hobbies encompass computing equipment, the mathematics of information, communications and invention. His success in the Shavian competition has stimulated wide interest in his proposed world alphabet (described in these pages) and a nonprofit Foundation for a Compatible and Consistent Alphabet has been set up and located at 333 N. Michigan Av. Malone resides with his wife and four children in Park Forest.

ØR FOIƎR. HO ORT ⊦N HEVEN. HALOED BI ⊥Δ NΔM.
⊥Δ K⊦ИDUM KUM. ⊥Δ WHL BI DUN. ΛN ƎRO ΛZ ⊦T ⊦Z ⊦N HEVEN.
G⊦V UC ⊥⊦C DΔ ØR DΔL⊦ BRED. ΛND FΛRG⊦V UC ØR TRECPΛCEZ.
ΛZ WI FΛRG⊦V ⊥OZ HO TRECPAC UGENCT UC. ΛND LID UC NOT
⊦NTO TEMPTΔSUN. BUT DIL⊦VƎR UC FRUM IV⊦L. FΛR ⊥ΔN ⊦Z ⊥U
K⊦ИDOM. ΛND ⊥U PØƎR. ΛND ⊥U GLOR⊦. FAREVƎR ΛND EVƎR. OMEN.

The familiar as it looks in alphabet proposed by John R. Malone. This, of course, is the Lord's Prayer, King James version.

TWO MORE SECTIONS INSIDE

Section Three, covering leisure activities, and Section Four, Feminine Angle and Your Home, are folded inside this section.

PROPOSED NEW ALPHABET

The Learning Time Is Reduced

Continued from Preceding Page

type machines are rapidly replacing typecasting machines and metal foundry type. Also, through use of photo platemaking and photo-offset printing, typewriter-like composing machines are being utilized more and more as the means of setting up large areas of printed material. *The cost of resetting or recomposing* the millions of words in contemporary English could be done in overseas areas to help build up the graphic arts industries in such countries as India, Africa and South America, just as Germany and America are helping to do in Taiwan and Japan at the present.

THE COST OF SUCH transliteration would be small indeed compared with the benefits that would accrue to the world by getting a common language.

And why not convert to a form of English which will help you gain access to the knowledge of other nations as well as to their contributions to knowledge. Much of the conversion during the next decade or so could be handled through the United Nations Educational Scientific and Cultural Organization or the International Documentary Center.

And for those who can already read English, it takes less than 45 minutes to retrain yourself *to read this—even an older dog can learn this* new trick.

WITHIN A FEW WEEKS it is possible to teach most 5- or 6-year-olds to write English with this alphabet. There are no rules or exceptions. It is ideal for teaching English to adults from non-English countries. Once confidence *and facility is attained, the problem of converting* to the older spelling forms is relatively easy, because of the compatibility feature, and memory devices built into this alphabet.

So far this alphabet has been tested in teaching children. It is also being used in a test class of Puerto Ricans. In each case the rate of learning is surprising. It is very much like teaching a person to count by means of Arabic numbers.

It may take a generation to get general concurrence for this type of alphabet for English, but given the long history of humanity, this is a relatively short time; and the economics possible with it are great. For instance, from 12 to 25 per cent fewer characters are needed to write a given piece of material in the simpler 40-character alphabet. The cost in reduced *learning time for youngsters should enable all* nations to upgrade their school systems, whether in America or abroad.

MANY TECHNICAL *developments, such* as machines for computing, accounting, check reading and for bibliographic listing and cataloging await adoption of this type of alphabet. So does the dictating typewriter, which takes the spoken word and types it out directly.

It is hardly needed to point out the commercial and political value in having a world speak English, as its common tongue. But the value is greatest to the poorer nations which would thereby have access to American-English techniques and scientific competence. The people of these countries could then live freely in our world, via the use of our methods and devices.

Now let's look at the alphabet itself a moment. For technical reasons all the letters have been designed with the same width, as typewriter letters generally are. To do this, some of the letters have been basically redesigned so that never again will I, l, or No. 1 be confused. There is no "lower case" or small letters as such. There is simply a flattened version of the same design. In this way needless configurations are eliminated.

FIRST WE HAVE added 11 vowel symbols to the A, E, I, O and U (and we have dispensed with the Y vowel usage). We have turned the 16 new vowels into five basic families, called the A, E, I, O and U families. The old letters are used to designate the "short" sounds of the letters as they are today: as cat, pet, bit, hot and but. Then there are added five *new "long vowels" for each of these as shown* in the following chart.

Each of these long vowels is characterized by a full width horizontal member Δ Ɪ Ʌ Ω Ʊ which helps you remember. Then there is an aw Λ, an e as in her Ǝ, a double o as in look Ⅎ, an ou as in couch Ø, and oy as in boy Ɑ and a yu sound as in you, or use Ш. Here is the way the five families look:

See 1983 Version

A: Λ Δ Λ
E: E Ǝ
I: Ɪ ⱡ
O: O Ō Ⅎ Ø Ɑ
U: U Ʊ Ш

In the consonant list the Cyrillic chay Ч has been added for the "ch" sound as in "chair"; the J has been broadened to fit our rectangle, "ng" as in sing has been turned into a single character Ͷ; s has become "sh" as in sure or shirt, while c has been made uniformly the "soft s" or "hiss" sound as in "cell" or "sin." The voiceless "th" of "thin" is represented by the Greek theta stylized to our purpose (neither the Romans or the later day Brooklyn citizens thought much of this sound). The voiced "th" as in "they" is represented by an upside down "T" ⊥. The modification of the W was made to keep it *from filling up with ink or dirt as it does* today on the typewriter or printing press. The Z has been crossed in the European fashion to distinguish it from a carelessly *made 2. A modification of the Cyrillic zh* sound character has been made for the most infrequently used sound in English, that found at the s in leisure, or the z in azure, and in many words of French-origin, such as rouge, beige, etc.

A FEW REMINDERS are needed: the G *is always hard as in "get"*; only K has the hard "c" sound now found in "crow." All buzzing "s" sounds as in business or glasses use z. Dropped entirely out are x and q *which can be replaced by combinations of* other letters.

Below are the 24 consonants and what they sound like:

See 1983 Version

B (b)	C (ss)	Ч (ch)	D (d)
F (f)	G (get)	H (h)	J (j)
K (k)	L (l)	M (m)	N (n)
Ͷ (ng)	P (p)	R (r)	S (sh)
T (t)	⊥ (th)	ϴ (th)	V (v)
W (w)	Y (yet)	Ƶ (z)	Ж (zh)

Now there is an added plus to this alphabet, besides such designed in features as being useful for computers and dictatable typewriters. It is sufficiently broad in phoneme representation so that it can be used for transcribing *Russian, Hebrew, Arabic, German, Italian and* Spanish phonetically.

With a few conventions or marks it can be used for French and Portuguese. The Romaji version of Japanese can go into it very easily and consistently with the present orthographic treatment.

OTHER PHONETIC alphabets have been proposed before, but this one is sufficiently comprehensive and practical for immediate use in primary schools at home, and in English training and technical schools at home and abroad. The technological conditions are ripening rapidly; the political, commercial and com*munication imperatives are clear* and demanding.

You can start writing this way tomorrow. You will find you can learn it easily, rapidly. *Write as you speak. English will never be the* same for you again — and lots easier to spell. You, American, will be considered among the most thoughtful people on earth—for you will have changed your ways so others can enjoy your movies, books, technics, riches and general cultural bounty and best of all, you can make your speech and language habits those of the world.

John R. Malone explains his 40-character alphabet and compatible number system. Lower case letters simply are smaller versions of the capital letters.

Figure 6b. Article from the Chicago Sunday Sun-Times discussing Unifon.

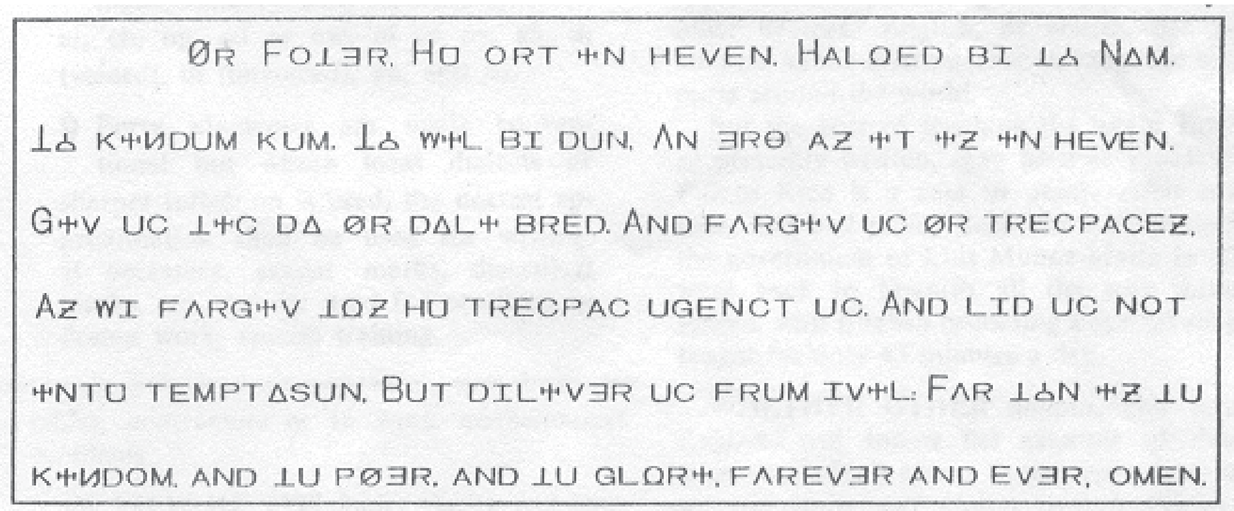


Figure 7. Example of an early version of Unifon (the alphabet as in Figure 1) set using upper- and lower-case. The alphabet given is:

AΔΛBCЧDEIƎFGHⱡᐃJKLMNͶOΩⱷØɑPRSTΘ⊥UƱШVWYƵX;

letters given here in red do not appear in the text.

Alisiz Advenȼʀz in Wundʀland

if mɨ hed *wɵd* gΩ ьru," ьΛt pɵr Alis, "it wɵd bɇ uv veri litul ʉs wiъɵt mɨ ꞩΩldʀz. Ω, hɵ Ɨ wiꞩ Ɨ kɵd ꞩut up lɨk u teluskΩp! Ɨ ьiиk Ɨ kɵd, if Ɨ Ωnli nu hɵ tu bigin." Fλr, ʉ sɇ, sΩ meni ɵt-uv-ъu-wΔ ьiиz had hapund lΔtli, ъat Alis had bigun tu ьiиk ъat veri fʉ ьiиz indɇd wʀ rɇuli imposubul.

Ъer sɇmd tu bɇ nΩ ʉs in wΔtiи bɨ ъu litul dλr, sΩ ꞩɇ went bak tu ъu tΔbul, haf hΩpiи ꞩɇ mɨt fɨnd unuъʀ kɇ on it, λr at eni rΔt u bɵk uv rʉlz fλr ꞩutiи pɇpul up lɨk teluskΩps: ъis tɨm ꞩɇ fɵnd u litul botul on it, ("hwiȼ sʀtunli wuz not hir bifλr," sed Alis,) and rɵnd ъu nek uv ъu botul wuz u pΔpʀ lΔbul, wiъ ъu wʀdz "DRIИK MɆ" bʉtufuli printid on it in lorj letʀz.

Figure 8. Example from Carroll [2012; in press], showing Unifon in a casing orthography. Carroll's English original likewise writes "DRINK ME" in all caps.

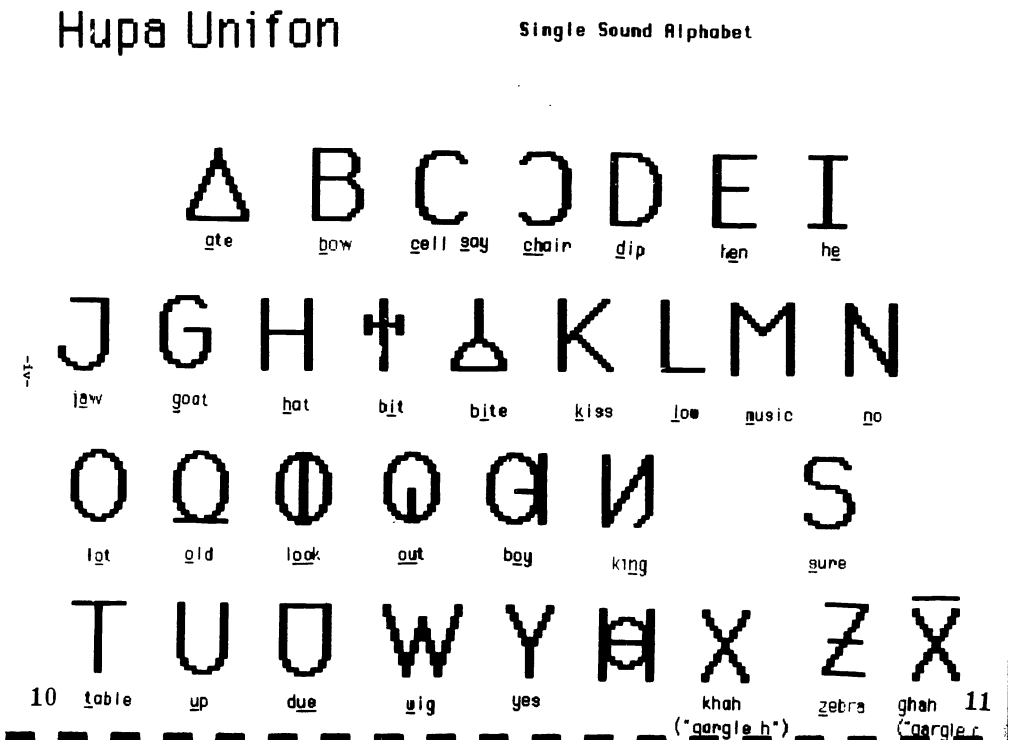


Figure 9. The Unifon alphabet for Hupe.

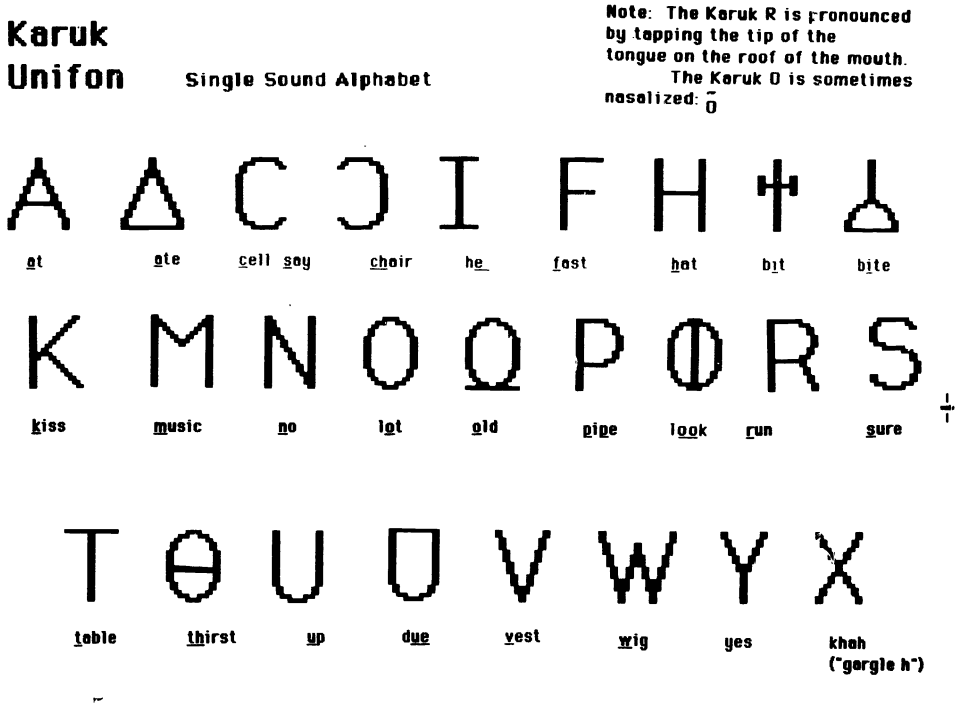


Figure 10. The Unifon alphabet for Karuk.

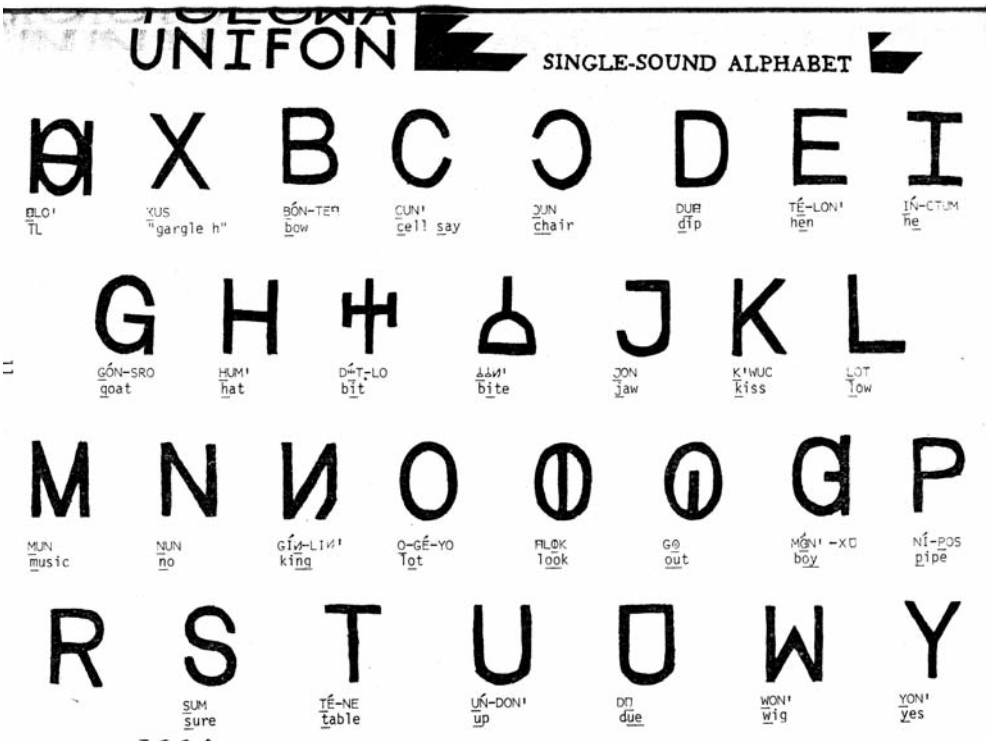


Figure 11. The Unifon alphabet for Tolowa.

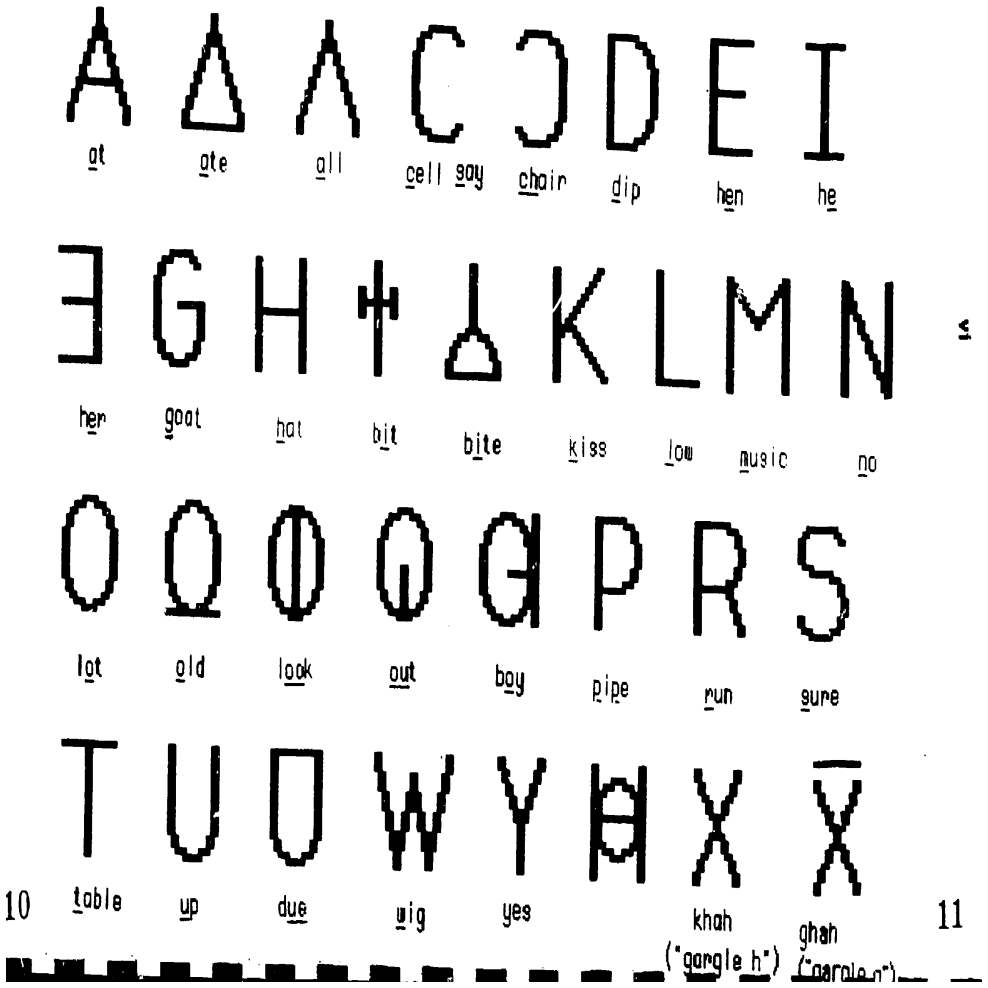


Figure 12. The Unifon alphabet for Yurok.

Figure 13. Evidently the source alphabet which was applied to different languages depending on their phonetic inventories.

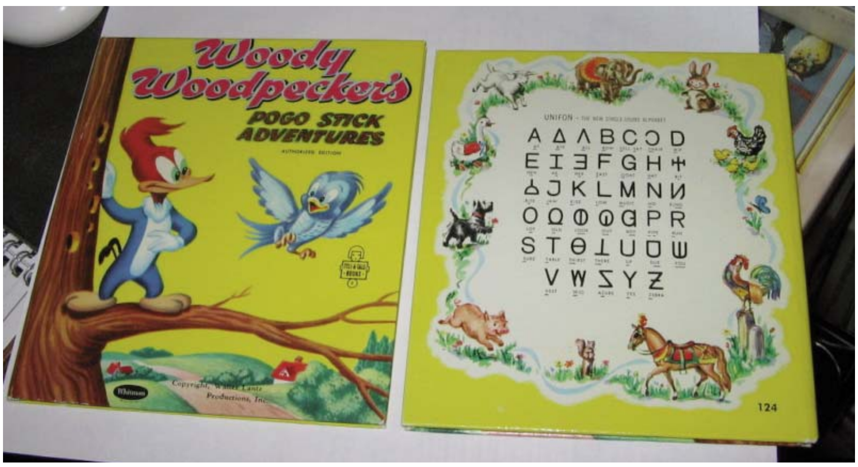


Figure 14. Example of a children's book published in Unifon in 1954. The alphabet given is:
AΔΛBCƆDEIƎFGHǂΔJKLMNИOΩΦΦGPRSTΘ⊥UŪШVWΣYƵ

A. Administrative

1. Title

Preliminary proposal to encode "Unifon" characters in the UCS.

2. Requester's name

UC Berkeley Script Encoding Initiative (Universal Scripts Project)
(Author) Michael Everson

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

Liaison contribution.

4. Submission date

2012-04-29

5. Requester's reference (if applicable)

6. Choose one of the following:

6a. This is a complete proposal

No.

6b. More information will be provided later

Yes.

B. Technical – General

1. Choose one of the following:

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

Not certain.

Proposed name of script

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

Not certain.

1c. Name of the existing block

2. Number of characters in proposal

Not certain.

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 A.

4. Is a repertoire including character names provided?

Yes.

4a. If YES, are the names in accordance with the "character naming guidelines"

Yes.

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

Yes.

5. Fonts related:

5a. Who will provide the appropriate computerized font to the Project Editor of 10646 for publishing the standard?

Michael Everson.

5b. Identify the party granting a license for use of the font by the editors (include address, e-mail, ftp-site, etc.):

Michael Everson, Fontlab and Fontographer.

6. References:

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. Special encoding issue: 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)?

No. Unicode character properties to be provided later.

8. Additional Information: 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.

See above.

C. Technical - Justification

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

No.

2. 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.

2a. If YES, available relevant documents

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

Linguists, teachers, educationists.

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

Rare.

4b. Reference

5a. Are the proposed characters in current use by the user community?
By some.
5b. If YES, where?
In the US.
6. After giving due considerations to the principles in the P&P document must the proposed characters be entirely in the BMP?
Not certain.
6a. If YES, is a rationale provided?
6b. If YES, reference
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?
Yes.
11b. If YES, is a rationale for such use provided?
No.
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?