JTC1/SC2/WG2 N4128
L2/11-280
2011-06-29

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

Doc Type: **Working Group Document**
Title: **Preliminary proposal for encoding the Moon script in the SMP of the UCS**
Source: **Michael Everson**
Status: **Individual Contribution**
Action: **For consideration by JTC1/SC2/WG2 and UTC**
Date: **2011-06-29**

1. Introduction.

The Moon System of Embossed Reading, also known as the Moon alphabet, Moon type, or Moon code) is a writing system for the blind which embossed symbols chiefly derived by applying simplifications to the shapes of the letters of the Latin alphabet. It was introduced in 1945 by William Moon (1818–1894), a blind Englishman from Brighton, East Sussex, who founded the Brighton Society for the Blind in 1860.

The Moon script is easily learned, particularly by those who have newly become blind and who find Braille daunting. There is evidence that some people who get used to reading by touch are able to transition from Moon script to Braille. Moon was adopted in Britain and in North America, though it was more popular in the 19th century and earlier 20th century than it is today. Nevertheless, a variety of printed literature exists in Moon script, and efforts are being made to adapt modern embossing technology

2. Processing. Short texts in Moon script are written from left to right, but longer embossed texts have boustrophedon directionality: at the end of a line there is a sort of parenthesis which guides the fingers to the next line, which the reader then follows. Glyphs are not mirrored in this style of boustrophedon writing.

3. Numerals. Moon script has three different systems of writing numbers. The first uses the letters O and I for 0 and 1, with the other digits being represented by diagonal strokes with a tick on one end. The second uses a unique NUMBER SIGN 𐏑 to indicate the numeric use of letters A through J as digits 1 through 0. When preceded by the letter L or P, this number sign indicates pounds and pence sterling. The third system of numbers was devised more recently to respond the the needs of new users, particularly those with learning difficulties, and uses shapes which are more iconic in terms of stroke count.

		0	1	2	3	4	5	6	7	8	9
System 1		𐏐	𐏒	𐏓	𐏔	𐏕	𐏖	𐏗	𐏘	𐏙	𐏚
System 2	𐏑	𐏛	𐏜	𐏝	𐏞	𐏟	𐏠	𐏡	𐏢	𐏣	𐏤
System 3		𐏥	𐏦	𐏧	𐏨	𐏩	𐏪	𐏫	𐏬	𐏭	𐏮

4. Punctuation. Moon script has its own punctuation marks. A single dot · is used for a COMMA, and a double dot ‥ is used for the FULL STOP. A triangle of three dots ∴ and inverted triangle of three dots ∵ are used for LEFT PARENTHESIS and RIGHT PARENTHESIS respectively. A QUESTION MARK ⸮ and EXCLAMATION MARK ! , and AMPERSAND ⅋ are used, as are an APOSTROPHE ' and a HYPHEN ‐ ; the last is a baseline character (because ― is LETTER T) and is also used as a decimal point (𐏓 𐏑 𐏔 ‐ 𐏝 𐏞 = £5.67).

The VERSE DIVIDER ⹀ is used to separate verses in poetry. The ABBREVIATION MARK ⁚ functions to show that letters have been left out, and a number of standard abbreviations have been specified. In addition to these (listed in the first column below), there are some one-letter abbreviations also specified:

⁚◝	:G	*-ing*	◡	B	*but*	◡	J	*just*	⁄	S	*some*
⁚—	:T	*-ment*	◜	C	*can*	◁	K	*know*	—	T	*that*
⁚◹	:N	*-tion*	◜·	CH	*child*	◺	L	*like*	◡	U	*under*
⁚⁄	:S	*-ness*	◞	D	*do*	◥	M	*more*	◡	V	*very*
⁚◞	:D	*-ound*	◸	E	*every*	◹	N	*not*	◠	W	*was*
⁚◸	:E	*-ence*	◜	F	*from*	○	O	*digit zero*	◠·	WH	*which*
⁚◺	:L	*-ful*	◝	G	*go*	◟	P	*people*	>	X	*it*
⁚◿	:Y	*-ity*	●	H	*have*	◿	Q	*quite*	◿	Y	*you*
			|	I	*digit one*	\	R	*rather*			

These can also be compounded as in ◜○◥⁚—⁚◝ COM:T:G *commenting*. The one-letter abbreviations can be compounded with the HYPHEN in longer words, as in ◁_◺◸◞◥◸ K-LEDGE *knowledge*. And both the HYPHEN and the ABBREVIATION MARK can be compounded, as in ◁_⁚◝ K-:G *knowing*, ◺_⁚⁄ L-:S *likeness*. In Swedish, the ABBREVIATION MARK ⁚ is used in front of a vowel as a diacritic, ⁚∧ :A *ä*. One "punctuation" character is the LOGOGRAM THE ≐, which is both used for the article *the* and to stand for *th* in general (like a kind of *þorn*):

≐ ≐⁚◝ ≐∧— ◜○∪◹—⁄ |⁄ ●◸∧≐◸\

Þ Þ:G ÞAT COUNTS IS HEAÞER

the thing that counts is heather

5. Character names. The names used for the characters here are based on the Latin letters from which the characters are devised; there are no other names used.

6. Repertoire extensions. Moon script was used with a certain amount of success for languages other than English. A variety of abbreviations and other conventions are used. Moon was also extended to cover other languages. In some of these extensions the original meanings of the characters are retained, but in others the shapes are used according to the shapes of the original alphabet they are intended to replace. In the examples in Figure 2, additional characters are used to extend the Latin alphabet snd to represent other alphabets. Transcriptions are given here:

English: ○∪\ ◜∧—●◸\ ∩●|◜● ∧\— |◹ ●◸∧∨◸◹.

OUR FATHER WHICH ART IN HEAVEN,

French: ◹○—\◸ ◟◸\◸ ◿∪| ◸⁄ ∧∪> ◜|◸∪>.

NOTRE PERE QUI ES AUX CIEUX,

German: ∪◹⁄◸\ ∨∧—◸\ |◹ ◞◸◥ ●|◥◥◸◺.

UNSER VATER IN DEM HIMMEL,

Dutch: ○◹◵◸ ∨∧◞◸\. ◞|◸ |◹ ◞◸ ●◸◥◸◺◸◹

ONZE VADER, DIE IN DE HEMELEN

Danish: ∨○\ ∨∧◞◸\. ◞∪ ⁄○◥ | ●|◥◺◸◹◸.

VOR VADER, DU SOM I HIMLENE,

Swedish: ᒥᐱᐠᒥᐦ ᐯᐱᐦ. ᐟᐅᒣ :ᐱᐦᐨ ᐁ ᐅᐁᒣᒪᐅᒣ.
FADER VÅR, SOM :ART I HIMLOM,

Russian: ᐅᑎᑌᒪ ∴ᐱᑎᔦ. ᑕᐟᑎᐁᐱ ∴ᐱ ∴ᒪᑌᒪᑕᐱᐳᔦ.
ОТЧЕ НАШЪ, СУЩІЙ НА НЕБЕСАХЪ,

Arabic: ᐁᐨᐁᑌᐁ ᐁᔾᒥᔭ ᑎᔭ ᐁᔾᐱᒥᐅᐟ ᔾᔭᐟᘔᒣ. ᐱ ᐁᐱᒥᔑ
ABANA ALḎY FY ALSMWAT LYTQD, S ASMK
ابانا الذي في السموات ليتقد, س اسمك

Armenian: ᑎᓀ <ᐨᔾᒥ ᐯᐟᒥ ᑎᒥ ᐟᒥᐱᐁIᑎ:ᐳᑎ ᐟᑌ.
O_ HAYR MER OR ERKIN:_N ES,
Ո_ ՀԱՅՐ ՄԵՐ ՈՐ ԵՐԿԻՆ:_Ն ԵՍ,

Greek: ᒥᑎ ᐱᐸᑕᒥ ᒥᑎ ·ᴏ ᒪᴏᖁᴏᐟ.
EN APXH HN ‘O ΛΟΓΟΣ,

Hindustani: ᐱᐁ ᴏᐱᒣ·ᐱᐸᒥ ᑌ·ᐱᔦ: ᔾᑐ ·ᐱᐟᒣ·ᐱᑎ ᔦᐱᐸ
AI NAM’ARE B’AP: JO ‘ASM’AN PAR

Ningbo: ᐱᴏ·ᒪᐱᴏ ᐱᴏ·ᐨᐁᐱ ᒪᔑ ᐨ·ᐁᑎ·ᘔᓀᐅᑎᖁ·ᖁᐅ:

AH-LAH AH-TIA LÆ T‘IN-Z^ONG-GO:

The issue here is that complete alphabets for the Cyrillic, Arabic, Greek, and Armenian have not yet been found. The chart below shows attested letters in red with conjectured letters left in black or (in the case of Armenian) given as hyphens.

Ʌ Ь ϒ Ↄ Г Ƶ Ґ ∴ I < L ℸ N > o ∠ ᒧ ᐟ — ∪ ᒥ ᑕ ᒧ O
A B Г Δ E Z H Θ I K Λ M N Ξ O Π P Σ T Y Ф X Ψ Ω

ا ب ت ث ج ح خ د ذ ر ز س ش ص ض ط ظ ع غ ف ق ك ل م ن ه و ي

— - - - ⁄ - - - - - ׀ - - - ᐱ < - - - V J ᑎ _ ᑎ - - - - ∪ - - ᒥ - - - - O -

Ա Բ Գ Դ Ե Զ Է Ը Թ Ժ Ի Լ Խ Ծ Կ Հ Ձ Ղ Ճ Մ Յ Ն Շ Ո Չ Պ Ջ Ռ Ս Վ Տ Ր Ց Ւ Փ Ք Օ Ֆ

Ʌ Ь ϒ Г Ↄ L ᑎ Z I Ʌ < J ℸ ∴ O — ᒧ ᑕ ᑎ ⁄ ᒥ > ∪ ᑌ ᑌ̇ ᑎ̇ ∠ Ƨ ↄ ⊥ ᴏ V
А Б В Г Д Е Ж З И Й К Л М Н О П Р С Т У Ф Х Ц Ч Ш Щ Ъ Ы Ь Э Ю Я

7. Sorting. The variant consonants are filed after their main consonants:

Ʌ *A* < Ⱥ *Á* < Ь *B* < ᑕ *C* < Ꞓ *CH* < Ↄ *D* < Г *E* < Ґ *É* < ᒥ *F* < ϒ *G* < ᴏ *H* <
I *I* < J *J* < < *K* < L *L* < ℸ *M* < ᑎ *N* < O *O* < ∠ *P* < ↄ *Q* < ᒧ *Q́* < ᒧ *R* <
⁄ *S* < — *T* < ∪ *U* < ᑌ *Ú* < V *V* < ᑎ *W* < ᑌ̇ *WH* < ᑎ̇ *Ẃ* < > *X* < ⊥ *Y* < Ƶ *Z*

8. Unicode Character Properties

To be supplied

9. Bibliography

Moon, William. 1877. *Light for the Blind: A history of the origin and success of Moon's System of Reading for the Blind*. London: Longmans & Co.

Moon Society. 1927. *The Moon Magazine*. Vol. 86, No. 5. London: National Institute for the Blind.

10. Figures.

Figure 1. Embossed chart of the Moon alphabet plus ampersand, from Moon 1877.

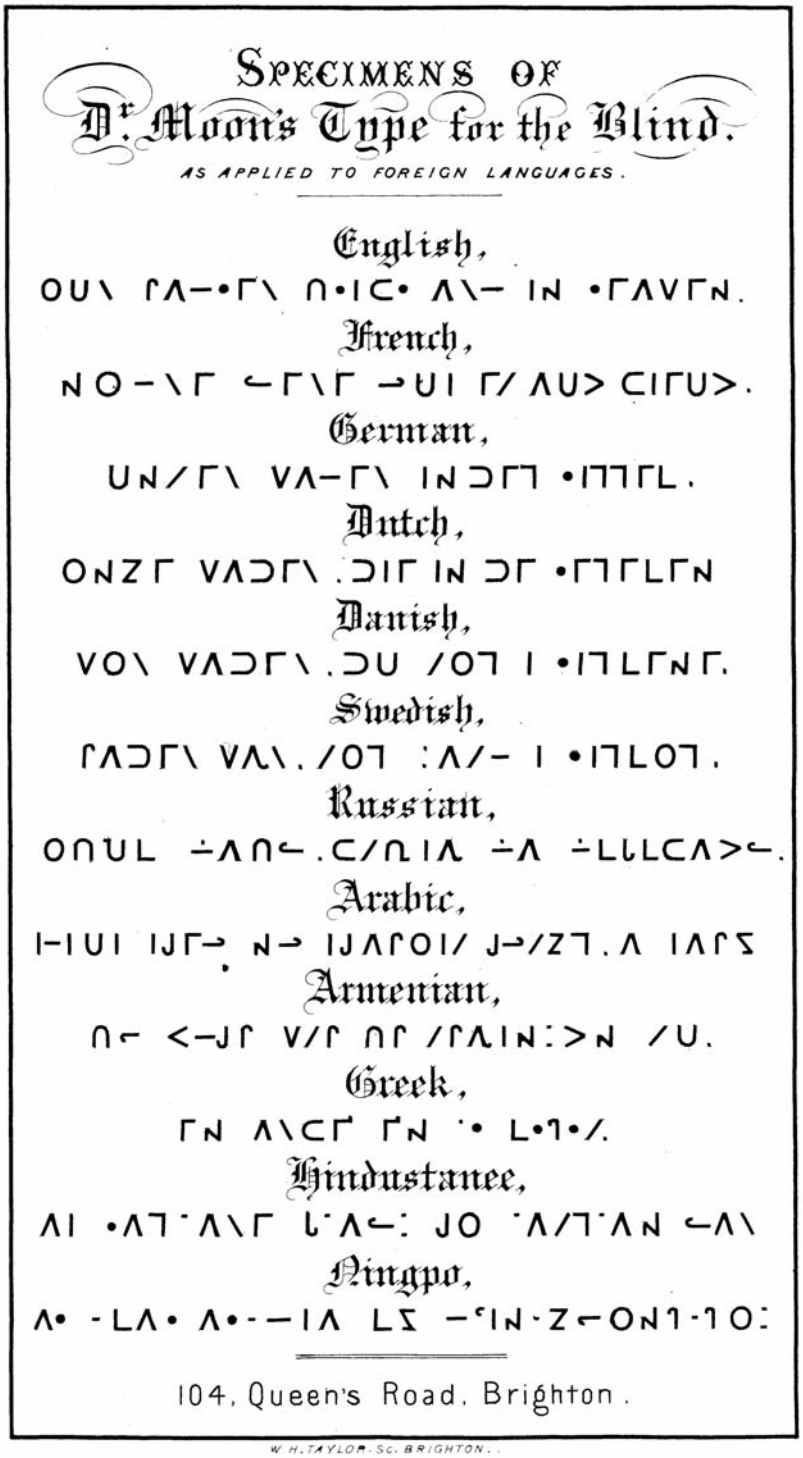
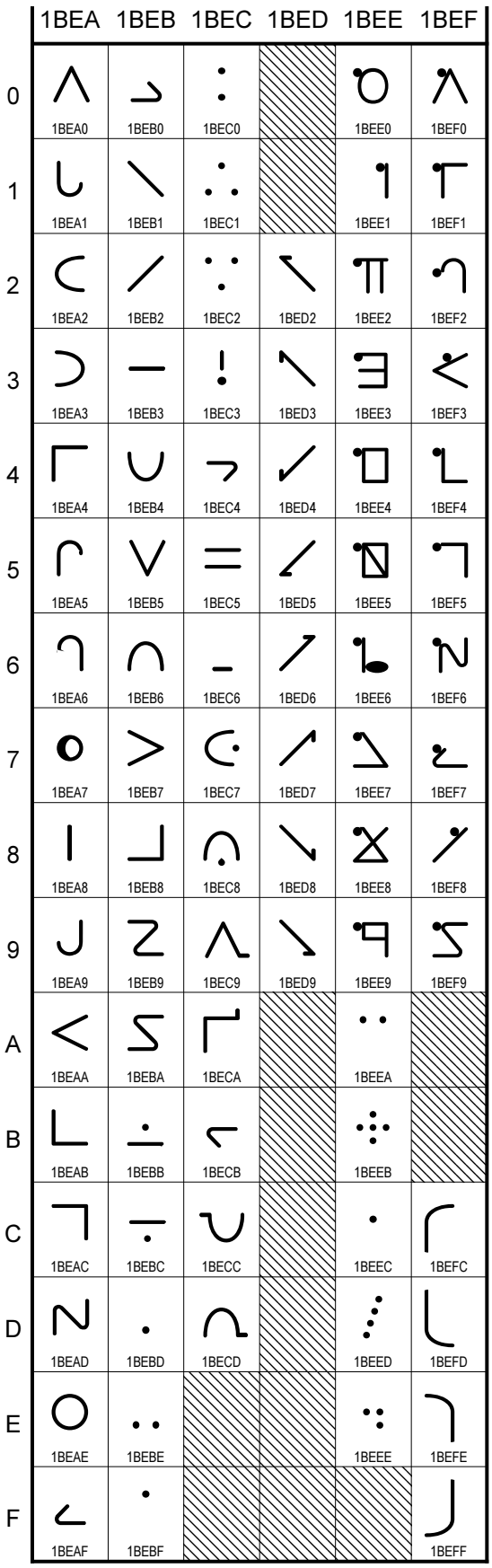


Figure 2. Chart of Moon extensions for foreign languages from Moon 1877.


1BEA 1BEB 1BEC 1BED 1BEE 1BEF
0 1BEA0 1BEB0 1BEC0 1BEE0 1BEF0
1 1BEA1 1BEB1 1BEC1 1BEE1 1BEF1
2 1BEA2 1BEB2 1BEC2 1BED2 1BEE2 1BEF2
3 1BEA3 1BEB3 1BEC3 1BED3 1BEE3 1BEF3
4 1BEA4 1BEB4 1BEC4 1BED4 1BEE4 1BEF4
5 1BEA5 1BEB5 1BEC5 1BED5 1BEE5 1BEF5
6 1BEA6 1BEB6 1BEC6 1BED6 1BEE6 1BEF6
7 1BEA7 1BEB7 1BEC7 1BED7 1BEE7 1BEF7
8 1BEA8 1BEB8 1BEC8 1BED8 1BEE8 1BEF8
9 1BEA9 1BEB9 1BEC9 1BED9 1BEE9 1BEF9
A 1BEAA 1BEBA 1BECA 1BEEA
B 1BEAB 1BEBB 1BECB 1BEEB
C 1BEAC 1BEBC 1BECC 1BEEC 1BEFC
D 1BEAD 1BEBD 1BECD 1BEED 1BEFD
E 1BEAE 1BEBE 1BEEE 1BEFE
F 1BEAF 1BEBF 1BEFF

Basic alphabet

1BEA0 MOON LETTER A
1BEA1 MOON LETTER B
1BEA2 MOON LETTER C
1BEA3 MOON LETTER D
1BEA4 MOON LETTER E
1BEA5 MOON LETTER F
1BEA6 MOON LETTER G
1BEA7 MOON LETTER H
1BEA8 MOON LETTER I
1BEA9 MOON LETTER J
1BEAA MOON LETTER K
1BEAB MOON LETTER L
1BEAC MOON LETTER M
1BEAD MOON LETTER N
1BEAE MOON LETTER O
1BEAF MOON LETTER P
1BEB0 MOON LETTER Q
1BEB1 MOON LETTER R
1BEB2 MOON LETTER S
1BEB3 MOON LETTER T
1BEB4 MOON LETTER U
1BEB5 MOON LETTER V
1BEB6 MOON LETTER W
1BEB7 MOON LETTER X
1BEB8 MOON LETTER Y
1BEB9 MOON LETTER Z

Punctuation

1BEBA MOON AMPERSAND
1BEBB MOON LOGOGRAM THE
• also used for th
1BEBC MOON NUMERAL SIGN
• precedes letters ABCDEFGHIJ used as digits 1234567890
1BEBD MOON COMMA
1BEBE MOON FULL STOP
1BEBF MOON APOSTROPHE
1BEC0 MOON ABBREVIATION SIGN
• used word-internally before letters to indicate abbreviation
1BEC1 MOON LEFT PARENTHESIS
1BEC2 MOON RIGHT PARENTHESIS
1BEC3 MOON EXCLAMATION MARK
1BEC4 MOON QUESTION MARK
1BEC5 MOON VERSE DIVIDER
• used to separate verses in poetry
1BEC6 MOON HYPHEN

Alphabetic extensions

1BEC7 MOON LETTER CH
1BEC8 MOON LETTER WH
1BEC9 MOON LETTER A WITH STROKE
• used for Swedish aa
• used for Cyrillic short i
1BECA MOON LETTER E WITH STROKE
• used for Greek eta
1BECB MOON LETTER TURNED Q
• used as a tone mark in Ningbo transcription
1BECC MOON LETTER U WITH STROKE
• used for Cyrillic che
1BECD MOON LETTER W WITH STROKE
• used for Cyrillic shcha

Old-style digits

1BED2 MOON NUMBER TWO
1BED3 MOON NUMBER THREE
1BED4 MOON NUMBER FOUR
1BED5 MOON NUMBER FIVE
1BED6 MOON NUMBER SIX
1BED7 MOON NUMBER SEVEN
1BED8 MOON NUMBER EIGHT
1BED9 MOON NUMBER NINE

StaffsMaths digits

1BEE0 MOON DIGIT ZERO
1BEE1 MOON DIGIT ONE
1BEE2 MOON DIGIT TWO
1BEE3 MOON DIGIT THREE
1BEE4 MOON DIGIT FOUR
1BEE5 MOON DIGIT FIVE
1BEE6 MOON DIGIT SIX
1BEE7 MOON DIGIT SEVEN
1BEE8 MOON DIGIT EIGHT
1BEE9 MOON DIGIT NINE

Additional punctuation (conjectured)

1BEEA MOON QUOTATION MARK
1BEEB MOON PLUS SIGN
1BEEC MOON MIDDLE DOT
1BEED MOON SOLIDUS
1BEEE MOON EQUALS SIGN

Additional letters (conjectured)

1BEF0 MOON LETTER DOTTED A
• addition
1BEF1 MOON LETTER DOTTED E
• equals
1BEF2 MOON LETTER DOTTED G
1BEF3 MOON LETTER DOTTED K
1BEF4 MOON LETTER DOTTED L
1BEF5 MOON LETTER DOTTED M
1BEF6 MOON LETTER DOTTED N
1BEF7 MOON LETTER DOTTED P
1BEF8 MOON LETTER DOTTED S
• subtraction
1BEF9 MOON LETTER DOTTED AMPERSAND

Guide brackets

1BEFC MOON TOP HALF LEFT BRACKET
1BEFD MOON BOTTOM HALF LEFT BRACKET
1BEFE MOON TOP HALF RIGHT BRACKET
1BEFF MOON BOTTOM HALF RIGHT BRACKET

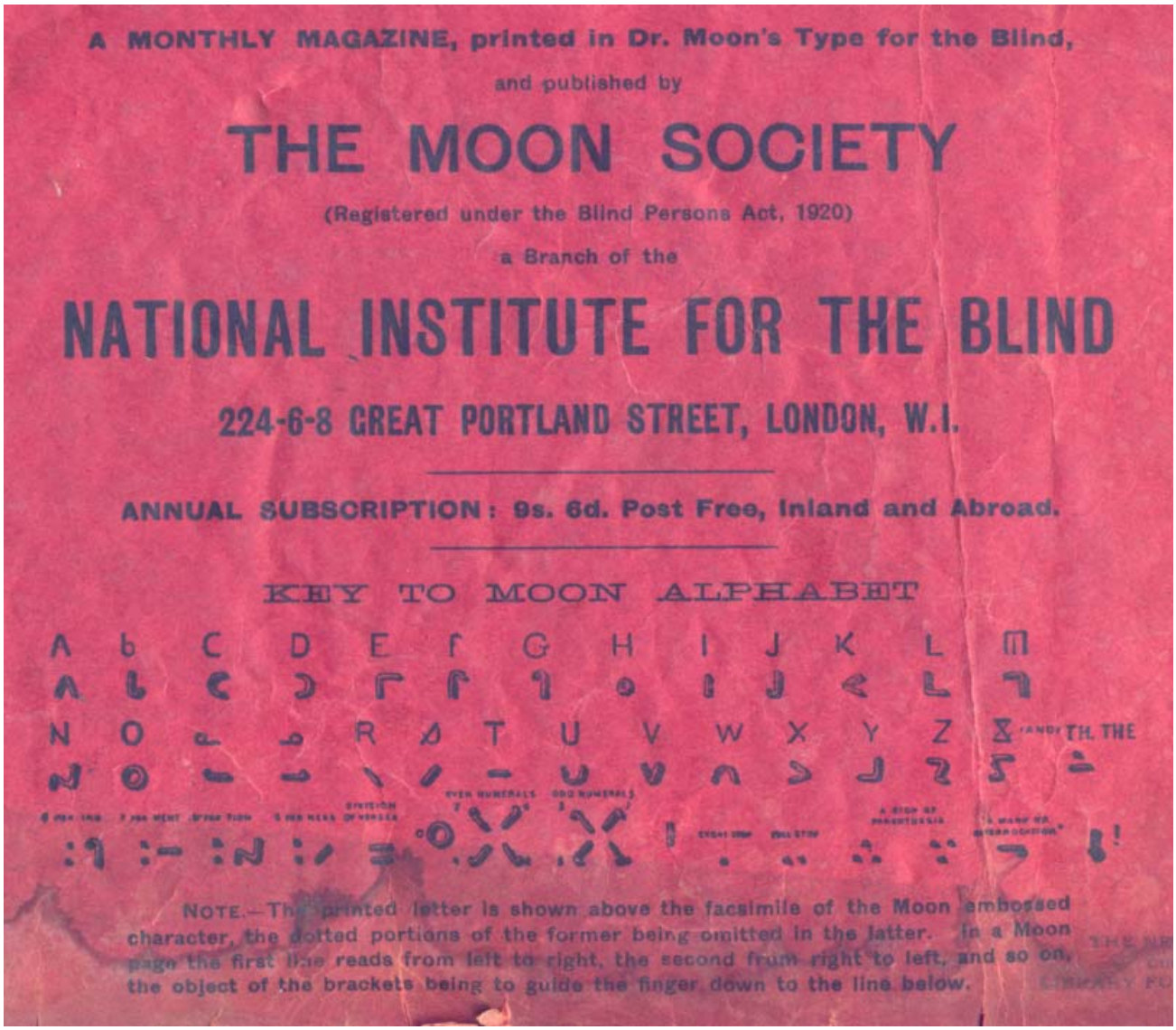


Figure 3. Chart of letters, punctuation, and numbers on the cover of the May 1927 edition of *The Moon Magazine*. This copy was previously held in the New York Public Libary' Library for the Blind.

Figure 4. The text of page 42 the May 1927 edition of *The Moon Magazine*. The text, part of a story titled “A Little Bit of Heather”, reads (with left-to-right directionality):

Figure 5. The Moon board on the printing floor of the Royal National Institue for the Blind. The size of the board is about 5 metres high

A. Administrative

1. Title

Proposal for encoding the Moon script in the SMP of the UCS

2. Requester's name

Michael Everson

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

Individual contribution.

4. Submission date

2011-06-29

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)

Yes.

1b. Proposed name of script

Moon.

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

40.

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 B-1.

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 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.

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 UAX #44 http://www.unicode.org/reports/tr44/ 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.

No.

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?

Dave Pawson.

2c. 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? Reference:

See above.

4a. The context of use for the proposed characters (type of use; common or rare)
Text for the blind.
4b. Reference
5a. Are the proposed characters in current use by the user community?
Yes.
5b. If YES, where? Reference:
Chiefly in the UK.
6a. After giving due considerations to the principles in the P&P document must the proposed characters be entirely in the BMP?
No.
6b. If YES, is a rationale provided?
6c. If YES, reference
7. Should the proposed characters be kept together in a contiguous range (rather than being scattered)?
Yes.
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?
Yes.
10b. If YES, is a rationale for its inclusion provided?
Yes.
10c. If YES, reference
The original shapes of the characters were derived from Latin.
11a. Does the proposal include use of combining characters and/or use of composite sequences?
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? If YES, reference