

DOC TYPE: Individual contribution
TITLE: **Proposal to encode Klingon in Unicode**
SOURCE: Mark Shoulson, Lieven Litaer
STATUS: Proposal
DISTRIBUTION: Worldwide

The Klingon alphabet (ᑭᑭᑭᑭ (ᑭᑭᑭᑭ), ISO-15924 code Piqd/293) has been the subject of some discussion in Unicode circles, beginning with (and perhaps before) its [initial proposal](#). Somewhat later, a review and demonstration of use was added to the Unicode Document Registry as [L2/16-329](#). A sampling of the images there is attached to this document, but we refer the reader to L2/16-329 for more extensive examples demonstrating the living usage of the script. Much of the text here derives from the original proposal (for which we thank Michael Everson, its author).

The encoding proposed derives from the Linux kernel source code (“Documentation/unicode.txt” by H. Peter Anvin, hpa@zytor.com), from which some of the description here has been taken. This proposal differs from the Linux registration by including two punctuation characters.

The standard character set used by the Klingon Language Institute is from a font which was created by the KLI, based on a set of ten characters designed for *Star Trek: The Motion Picture* by the Astra Image Corporation (which CBS/Paramount sometimes uses for decoration, and has never connected with ᑭᑭᑭᑭ ᑭᑭᑭᑭ, the Klingon language).

Klingon has an alphabet of 26 characters (ᑭᑭᑭᑭᑭᑭ), a positional numeric writing system with 10 digits, and is written left-to-right, top-to-bottom. Punctuation in the form of triangles for commas and full stops is attested in the examples. The standard Okrand transliteration for Klingon runs as follows:

a b ch D e gh H I j l m n ng o p q Q r S t tlh u v w y '

The vowel letters are named ᑭᑭᑭᑭ, ᑭᑭᑭᑭ, ᑭᑭᑭᑭ, ᑭᑭᑭᑭ and ᑭᑭᑭᑭ¹, and the consonant letters (except for ᑭ) are named by the syllable formed by appending «-ᑭᑭᑭᑭ» (“-ay”, pronounced like English “eye”) to the consonant. The letter ᑭ (pronounced as a glottal stop) is exceptional, being named ᑭᑭᑭᑭᑭᑭ («qaghwiᑭ» literally, “interrupter”). We have dropped the ' from the letter names in the transliteration for the character list, and spelled «Q» as “QH”.

This encoding has been endorsed by the Klingon Language Institute (<http://www.kli.org/>).

¹ «'at», «'et», «'it», «'ot» and «'ut»

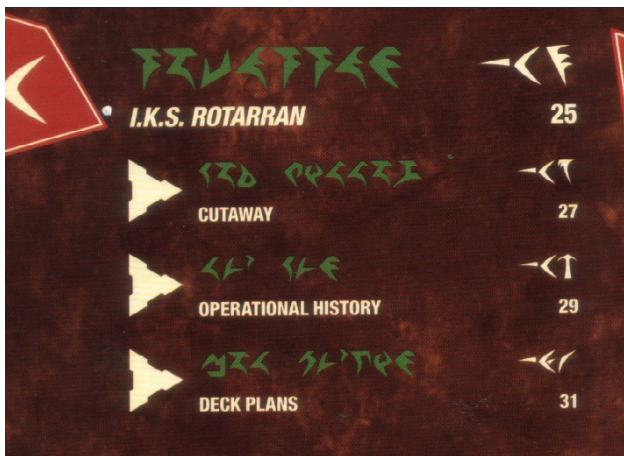
U+1CDD0 KLINGON LETTER AT;Lo;0;L;;;;;N;;;;;
U+1CDD1 KLINGON LETTER BAY;Lo;0;L;;;;;N;;;;;
U+1CDD2 KLINGON LETTER CHAY;Lo;0;L;;;;;N;;;;;
U+1CDD3 KLINGON LETTER DAY;Lo;0;L;;;;;N;;;;;
U+1CDD4 KLINGON LETTER ET;Lo;0;L;;;;;N;;;;;
U+1CDD5 KLINGON LETTER GHAY;Lo;0;L;;;;;N;;;;;
U+1CDD6 KLINGON LETTER HAY;Lo;0;L;;;;;N;;;;;
U+1CDD7 KLINGON LETTER IT;Lo;0;L;;;;;N;;;;;
U+1CDD8 KLINGON LETTER JAY;Lo;0;L;;;;;N;;;;;
U+1CDD9 KLINGON LETTER LAY;Lo;0;L;;;;;N;;;;;
U+1CDDA KLINGON LETTER MAY;Lo;0;L;;;;;N;;;;;
U+1Cddb KLINGON LETTER NAY;Lo;0;L;;;;;N;;;;;
U+1CDDC KLINGON LETTER NGAY;Lo;0;L;;;;;N;;;;;
U+1CDDD KLINGON LETTER OT;Lo;0;L;;;;;N;;;;;
U+1CDD E KLINGON LETTER PAY;Lo;0;L;;;;;N;;;;;
U+1CDDF KLINGON LETTER QAY;Lo;0;L;;;;;N;;;;;
U+1CDE0 KLINGON LETTER QHAY;Lo;0;L;;;;;N;;;;;
U+1CDE1 KLINGON LETTER RAY;Lo;0;L;;;;;N;;;;;
U+1CDE2 KLINGON LETTER SAY;Lo;0;L;;;;;N;;;;;
U+1CDE3 KLINGON LETTER TAY;Lo;0;L;;;;;N;;;;;
U+1CDE4 KLINGON LETTER TLHAY;Lo;0;L;;;;;N;;;;;
U+1CDE5 KLINGON LETTER UT;Lo;0;L;;;;;N;;;;;
U+1CDE6 KLINGON LETTER VAY;Lo;0;L;;;;;N;;;;;
U+1CDE7 KLINGON LETTER WAY;Lo;0;L;;;;;N;;;;;
U+1CDE8 KLINGON LETTER YAY;Lo;0;L;;;;;N;;;;;
U+1CDE9 KLINGON LETTER QAGHWI;Lo;0;L;;;;;N;;;;;
U+1CDF0 KLINGON DIGIT ZERO;Nd;0;L;;0;0;0;N;;;;;
U+1CDF1 KLINGON DIGIT ONE;Nd;0;L;;1;1;1;N;;;;;
U+1CDF2 KLINGON DIGIT TWO;Nd;0;L;;2;2;2;N;;;;;
U+1CDF3 KLINGON DIGIT THREE;Nd;0;L;;3;3;3;N;;;;;
U+1CDF4 KLINGON DIGIT FOUR;Nd;0;L;;4;4;4;N;;;;;
U+1CDF5 KLINGON DIGIT FIVE;Nd;0;L;;5;5;5;N;;;;;
U+1CDF6 KLINGON DIGIT SIX;Nd;0;L;;6;6;6;N;;;;;
U+1CDF7 KLINGON DIGIT SEVEN;Nd;0;L;;7;7;7;N;;;;;
U+1CDF8 KLINGON DIGIT EIGHT;Nd;0;L;;8;8;8;N;;;;;
U+1CDF9 KLINGON DIGIT NINE;Nd;0;L;;9;9;9;N;;;;;
U+1CDFD KLINGON COMMA;Po;0;L;;;;;N;;;;;
U+1CDFE KLINGON FULL STOP;Po;0;L;;;;;N;;;;;

U+1CDD0 KLINGON LETTER AT
 U+1CDD1 KLINGON LETTER BAY
 U+1CDD2 KLINGON LETTER CHAY
 U+1CDD3 KLINGON LETTER DAY
 U+1CDD4 KLINGON LETTER ET
 U+1CDD5 KLINGON LETTER GHAY
 U+1CDD6 KLINGON LETTER HAY
 U+1CDD7 KLINGON LETTER IT
 U+1CDD8 KLINGON LETTER JAY
 U+1CDD9 KLINGON LETTER LAY
 U+1CDDA KLINGON LETTER MAY
 U+1CDDB KLINGON LETTER NAY
 U+1CDDC KLINGON LETTER NGAY
 U+1CDDD KLINGON LETTER OT
 U+1CDDE KLINGON LETTER PAY
 U+1CDDF KLINGON LETTER QAY
 U+1CDE0 KLINGON LETTER QHAY
 U+1CDE1 KLINGON LETTER RAY
 U+1CDE2 KLINGON LETTER SAY
 U+1CDE3 KLINGON LETTER TAY
 U+1CDE4 KLINGON LETTER TLHAY
 U+1CDE5 KLINGON LETTER UT
 U+1CDE6 KLINGON LETTER VAY
 U+1CDE7 KLINGON LETTER WAY
 U+1CDE8 KLINGON LETTER YAY
 U+1CDE9 KLINGON LETTER QAGHWI
 U+1CDEA (This position shall not be used)
 U+1CDEB (This position shall not be used)
 U+1CDEB (This position shall not be used)
 U+1CDED (This position shall not be used)
 U+1CDEE (This position shall not be used)
 U+1CDEF (This position shall not be used)
 U+1CDF0 KLINGON DIGIT ZERO
 U+1CDF1 KLINGON DIGIT ONE
 U+1CDF2 KLINGON DIGIT TWO
 U+1CDF3 KLINGON DIGIT THREE
 U+1CDF4 KLINGON DIGIT FOUR
 U+1CDF5 KLINGON DIGIT FIVE
 U+1CDF6 KLINGON DIGIT SIX
 U+1CDF7 KLINGON DIGIT SEVEN
 U+1CDF8 KLINGON DIGIT EIGHT
 U+1CDF9 KLINGON DIGIT NINE
 U+1CDFA (This position shall not be used)
 U+1CDFB (This position shall not be used)
 U+1CDFC (This position shall not be used)
 U+1CDFD KLINGON COMMA
 U+1CDFE KLINGON FULL STOP
 U+1CFFF (This position shall not be used)

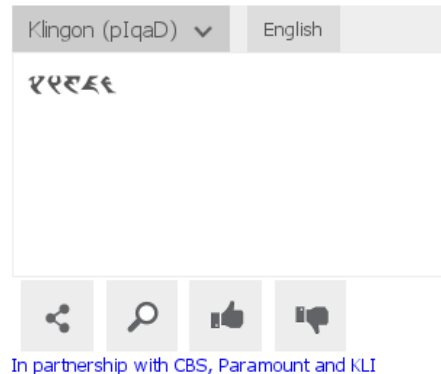
	1CDD	1CDE	1CDF
0			
1			
2			
3			
4			
5			
6			
7			
8			
9			
A			
B			
C			
D			
E			
F			



Comic book



Book glossary



In partnership with CBS, Paramount and KLI

Bing Translate



Monopoly board

**ISO/IEC JTC 1/SC 2/WG 2
PROPOSAL SUMMARY FORM TO ACCOMPANY SUBMISSIONS
FOR ADDITIONS TO THE REPERTOIRE OF ISO/IEC 10646².**

Please fill all the sections A, B and C below.

Please read Principles and Procedures Document (P & P) from <http://std.dkuug.dk/JTC1/SC2/WG2/docs/principles.html> for guidelines and details before filling this form.

Please ensure you are using the latest Form from <http://std.dkuug.dk/JTC1/SC2/WG2/docs/summaryform.html>.

See also <http://std.dkuug.dk/JTC1/SC2/WG2/docs/roadmaps.html> for latest Roadmaps.

A. Administrative

1. Title:	<i>Proposal to encode Klingon in Unicode</i>
2. Requester's name:	<i>Mark Shoulson, Lieven Litaer</i>
3. Requester type (Member body/Liaison/Individual contribution):	<i>Individual Request</i>
4. Submission date:	
5. Requester's reference (if applicable):	
6. Choose one of the following:	
This is a complete proposal:	<i>Yes</i>
(or) More information will be provided later:	<i>No</i>

B. Technical – General

1. Choose one of the following:	
a. This proposal is for a new script (set of characters):	<i>Yes</i>
Proposed name of script:	<i>Klingon</i>
b. The proposal is for addition of character(s) to an existing block:	<i>No</i>
Name of the existing block:	
2. Number of characters in proposal:	<i>37</i>
3. Proposed category (select one from below - see section 2.2 of P&P document):	
A-Contemporary <input type="checkbox"/> B.1-Specialized (small collection) <input checked="" type="checkbox"/> B.2-Specialized (large collection) <input type="checkbox"/>	
C-Major extinct <input type="checkbox"/> D-Attested extinct <input type="checkbox"/> E-Minor extinct <input type="checkbox"/>	
F-Archaic Hieroglyphic or Ideographic <input type="checkbox"/> G-Obscure or questionable usage symbols <input type="checkbox"/>	
4. Is a repertoire including character names provided?	<i>Yes</i>
a. If YES, are the names in accordance with the "character naming guidelines" in Annex L of P&P document?	<i>Y</i>
b. Are the character shapes attached in a legible form suitable for review?	<i>Yes</i>
5. Fonts related:	
a. Who will provide the appropriate computerized font to the Project Editor of 10646 for publishing the standard?	<i>Michael Everson, Evertyp / Mark Shoulson, KLI</i>
b. Identify the party granting a license for use of the font by the editors (include address, e-mail, ftp-site, etc.):	
6. References:	
a. Are references (to other character sets, dictionaries, descriptive texts etc.) provided?	<i>Yes</i>
b. Are published examples of use (such as samples from newspapers, magazines, or other sources) of proposed characters attached?	<i>Yes</i>
7. Special encoding issues:	
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)?	<i>No</i>

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. 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/reports/tr44/>) and associated Unicode Technical Reports for information needed for consideration by the Unicode Technical Committee for inclusion in the Unicode Standard.

C. Technical - Justification

1. Has this proposal for addition of character(s) been submitted before? If YES explain	<i>N1643; L2/16-329 cites further examples of usage</i>	<i>Yes</i>
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.)? If YES, with whom? If YES, available relevant documents:	<i>Klingon Language Institute</i>	<i>Yes</i>
3. Information on the user community for the proposed characters (for example: size, demographics, information technology use, or publishing use) is included? Reference:		<i>Yes</i>
4. The context of use for the proposed characters (type of use; common or rare) Reference:		<i>Rare. hobbyist</i>
5. Are the proposed characters in current use by the user community? If YES, where? Reference:	<i>By scholars and enthusiasts</i>	<i>Yes</i>
6. After giving due considerations to the principles in the P&P document must the proposed characters be entirely in the BMP? If YES, is a rationale provided? If YES, reference:		<i>No</i>
7. Should the proposed characters be kept together in a contiguous range (rather than being scattered)?		<i>Yes</i>
8. Can any of the proposed characters be considered a presentation form of an existing character or character sequence? If YES, is a rationale for its inclusion provided? If YES, reference:		<i>No</i>
9. Can any of the proposed characters be encoded using a composed character sequence of either existing characters or other proposed characters? If YES, is a rationale for its inclusion provided? If YES, reference:		<i>No</i>
10. Can any of the proposed character(s) be considered to be similar (in appearance or function) to, or could be confused with, an existing character? If YES, is a rationale for its inclusion provided? If YES, reference:		<i>No</i>
11. Does the proposal include use of combining characters and/or use of composite sequences? If YES, is a rationale for such use provided? If YES, reference: Is a list of composite sequences and their corresponding glyph images (graphic symbols) provided? If YES, reference:		<i>No</i>
12. Does the proposal contain characters with any special properties such as control function or similar semantics? If YES, describe in detail (include attachment if necessary)		<i>No</i>
13. Does the proposal contain any Ideographic compatibility characters? If YES, are the equivalent corresponding unified ideographic characters identified? If YES, reference:		<i>No</i>