

**Draft Proposal for formal registration of TSCII bilingual encoding  
with IETF/IANA as per RFC 2978 guidelines**

To: [ietf-charsets@iana.org](mailto:ietf-charsets@iana.org)

Subject: Registration of new charset

Character set name:

TSCII

(TAMIL SCRIPT CODE FOR INFORMATION INTERCHANGE)

Character set aliases:

None

Suitability for use in MIME text:

YES

Published Specifications:

[www.tscii.org/tscii/spec.html](http://www.tscii.org/tscii/spec.html)

ISO 10646 Equivalency Table:

<http://www.unicode.org/notes/tn15/>

Intended usage:

COMMON

Additional Information:

Tamil is one of the main Indian languages (Dravidian in Origin) currently spoken by over 70 million people worldwide. TSCII (Tamil Script Code for Information Interchange) is a bilingual 8-bit glyph-based encoding scheme (Roman and Tamil) to deal with Tamil materials on computers and for Information Interchange across platforms using different protocols and document formats. The TSCII scheme was collectively worked out through Net-based discussions in 1998. TSCII is modelled on the ISO-8859-XX scheme with standard plain ASCII set filling the 7-bit part

and a set of Tamil character glyphs filling the 8-bit part.

The TSCII scheme has been widely in use for over 5 years in all three popular computer platforms (Windows, Macintosh and Unix/Linux). In addition to millions of home-users (particularly in India, Singapore, Malaysia, Sri Lanka Western Europe and North America), TSCII encoding is used widely in Net-based mailing lists, newspapers and ezines on-line, digital library etc. Legacy data in TSCII format generated during the last 5 years is quite substantial and is growing constantly.

TSCII as an established language encoding is already recognized by major IT players like the Unicode Consortium, Microsoft, Apple, Oracle and Sun Microsystems. With OS-level support for Tamil in Microsoft Windows 2000 and later OS releases and very recently in Apple's Mac OS X 10.4 (Tiger) release, Tamil Diaspora has started to use Unicode already. The Purpose of this formal registration with IETF is to facilitate migration of the vast amounts of legacy data in TSCII and multitude of users.

Person(s) & email address to contact for further information:

TSCII USER GROUP represented by

Kalyanasundaram, Kuppuswamy (Switzerland)

[kalyan.geo@yahoo.com](mailto:kalyan.geo@yahoo.com)

Manivannan, Mani (USA)

[mmanivannan@gmail.com](mailto:mmanivannan@gmail.com)

Nedumaran, Muthu (Malaysia)

[muthu@murasu.com](mailto:muthu@murasu.com)

Kaviarasan, S (USA)

[kaviarasan@yahoo.com](mailto:kaviarasan@yahoo.com)

Paul, Ravindran K (Malaysia)

[ravi@thunaivan.com](mailto:ravi@thunaivan.com)

Doddannan, Sivaraj (India)

[sivaraj@theni.net](mailto:sivaraj@theni.net)

RM. Krishnan (India)

[poo@giasmd01.vsnl.net.in](mailto:poo@giasmd01.vsnl.net.in)

Kumar Mallikarjunan (USA)

[kumar@exchange.vt.edu](mailto:kumar@exchange.vt.edu)

Sinnathurai Srivas (UK)

[SiSrivias@hotmail.com](mailto:SiSrivias@hotmail.com)

## TAMIL SCRIPT CODE FOR INFORMATION INTERCHANGE (TSCII)

### Glyph/Character Listing and ISO 10646 Mapping Table

Format: Three tab-separated columns

Column #1 is character/glyph graphic representation

Column #2 is the TSCII code position (in hex)

Column #3 is the TSCII character name

ISO 10646 Mapping table can be obtained from

<http://www.unicode.org/notes/tn15>

## Character HEX Character Name

<control>	00	NULL
<control>	01	START OF HEADING
<control>	02	START OF TEXT
<control>	03	END OF TEXT
<control>	04	END OF TRANSMISSION
<control>	05	ENQUIRY
<control>	06	ACKNOWLEDGE
<control>	07	BELL
<control>	08	BACKSPACE
<control>	09	HORIZONTAL TABULATION
<control>	0A	LINE FEED
<control>	0B	VERTICAL TABULATION
<control>	0C	FORM FEED
<control>	0D	CARRIAGE RETURN
<control>	0E	SHIFT OUT
<control>	0F	SHIFT IN
<control>	10	DATA LINE ESCAPE
<control>	11	DEVICE CONTROL ONE
<control>	12	DEVICE CONTROL TWO
<control>	13	DEVICE CONTROL THREE
<control>	14	DEVICE CONTROL FOUR
<control>	15	NEGATIVE ACKNOWLEDGE
<control>	16	SYNCHRONOUS IDLE
<control>	17	END OF TRANSMISSION BLOCK
<control>	18	CANCEL
<control>	19	END OF MEDIUM
<control>	1A	SUBSTITUTE
<control>	1B	ESCAPE

<control>	1C	FILE SEPARATOR
<control>	1D	GROUP SEPARATOR
<control>	1E	RECORD SEPARATOR
<control>	1F	UNIT SEPARATOR

	20	SPACE
!	21	EXCLAMATION MARK
"	22	QUOTATION MARK
#	23	NUMBER SIGN
\$	24	DOLLAR SIGN
%	25	PERCENT SIGN
&	26	AMPERSAND
'	27	APOSTROPHE
(	28	LEFT PARENTHESIS
)	29	RIGHT PARENTHESIS
*	2A	ASTERISK
+	2B	PLUS SIGN
,	2C	COMMA
-	2D	HYPHEN MINUS
.	2E	FULL STOP
/	2F	SOLIDUS
0	30	DIGIT ZERO
1	31	DIGIT ONE
2	32	DIGIT TWO
3	33	DIGIT THREE
4	34	DIGIT FOUR
5	35	DIGIT FIVE
6	36	DIGIT SIX
7	37	DIGIT SEVEN
8	38	DIGIT EIGHT
9	39	DIGIT NINE
:	3A	COLON
;	3B	SEMICOLON
<	3C	LESS-THAN SIGN
=	3D	EQUALS SIGN

>	3E	GREATER-THAN SIGN
?	3F	QUESTION MARK
@	40	COMMERCIAL AT
A	41	LATIN CAPITAL LETTER A
B	42	LATIN CAPITAL LETTER B
C	43	LATIN CAPITAL LETTER C
D	44	LATIN CAPITAL LETTER D
E	45	LATIN CAPITAL LETTER E
F	46	LATIN CAPITAL LETTER F
G	47	LATIN CAPITAL LETTER G
H	48	LATIN CAPITAL LETTER H
I	49	LATIN CAPITAL LETTER I
J	4A	LATIN CAPITAL LETTER J
K	4B	LATIN CAPITAL LETTER K
L	4C	LATIN CAPITAL LETTER L
M	4D	LATIN CAPITAL LETTER M
N	4E	LATIN CAPITAL LETTER N
O	4F	LATIN CAPITAL LETTER O
P	50	LATIN CAPITAL LETTER P
Q	51	LATIN CAPITAL LETTER Q
R	52	LATIN CAPITAL LETTER R
S	53	LATIN CAPITAL LETTER S
T	54	LATIN CAPITAL LETTER T
U	55	LATIN CAPITAL LETTER U
V	56	LATIN CAPITAL LETTER V
W	57	LATIN CAPITAL LETTER W
X	58	LATIN CAPITAL LETTER X
Y	59	LATIN CAPITAL LETTER Y
Z	5A	LATIN CAPITAL LETTER Z
[	5B	LEFT SQUARE BRACKET



\	5C	REVERSE SOLIDUS
]	5D	RIGHT SQUARE BRACKET
^	5E	CIRCUMFLEX ACCENT
_	5F	LOW LINE
`	60	GRAVE ACCENT
a	61	LATIN SMALL LETTER A
b	62	LATIN SMALL LETTER B
c	63	LATIN SMALL LETTER C
d	64	LATIN SMALL LETTER D
e	65	LATIN SMALL LETTER E
f	66	LATIN SMALL LETTER F
g	67	LATIN SMALL LETTER G
h	68	LATIN SMALL LETTER H
i	69	LATIN SMALL LETTER I
j	6A	LATIN SMALL LETTER J
k	6B	LATIN SMALL LETTER K
l	6C	LATIN SMALL LETTER L
m	6D	LATIN SMALL LETTER M
n	6E	LATIN SMALL LETTER N
o	6F	LATIN SMALL LETTER O
p	70	LATIN SMALL LETTER P
q	71	LATIN SMALL LETTER Q
r	72	LATIN SMALL LETTER R
s	73	LATIN SMALL LETTER S
t	74	LATIN SMALL LETTER T
u	75	LATIN SMALL LETTER U
v	76	LATIN SMALL LETTER V
w	77	LATIN SMALL LETTER W
x	78	LATIN SMALL LETTER X
y	79	LATIN SMALL LETTER Y

Z	7A	LATIN SMALL LETTER Z
{	7B	LEFT CURLY BRACKET
	7C	VERTICAL LINE
}	7D	RIGHT CURLY BRACKET
~	7E	TILDE
	7F	DELETE
o	80	TAMIL DIGIT CUZHI = Tamil digit zero
க	81	TAMIL DIGIT ONRRU = Tamil digit one
ஸ்ரீ	82	TAMIL GRANTHA LETTER SRI = Tamil letter sri
ஜ	83	TAMIL GRANTHA LETTER JA = Tamil letter ja
ஷ	84	TAMIL GRANTHA LETTER SSA = Tamil letter ssa
ஸ	85	TAMIL GRANTHA LETTER SA = Tamil letter sa
ஹ	86	TAMIL GRANTHA LETTER HA = Tamil letter ha
க்ஷ	87	TAMIL GRANTHA LETTER KSHA = Tamil letter ksha
ஜ	88	TAMIL GRANTHA LETTER J = Tamil letter j
ஷ	89	TAMIL GRANTHA LETTER SS = Tamil letter ss
ஸ்	8A	TAMIL GRANTHA LETTER S = Tamil letter s
ஹ	8B	TAMIL GRANTHA LETTER H = Tamil letter h

க்ஷ	8C	TAMIL GRANTHA LETTER KSH = Tamil letter ksh
௨	8D	TAMIL DIGIT IRANNNTU = Tamil digit two
௩	8E	TAMIL DIGIT MUUNNRRU = Tamil digit three
௪	8F	TAMIL DIGIT NAANNKU = Tamil digit four
௫	90	TAMIL DIGIT AINTHU = Tamil digit five
‘	91	LEFT SINGLE QUOTATION MARK
’	92	RIGHT SINGLE QUOTATION MARK
“	93	LEFT DOUBLE QUOTATION MARK
”	94	RIGHT DOUBLE QUOTATION MARK
௬	95	TAMIL DIGIT AARRU = Tamil digit six
௭	96	TAMIL DIGIT EEZHU = Tamil digit seven
௮	97	TAMIL DIGIT ETTU = Tamil digit eight
௯	98	TAMIL DIGIT ONPATHU = Tamil digit nine
ங்	99	TAMIL LETTER NGAKARA UKARAM = Tamil letter ngu
ஞ்	9A	TAMIL LETTER NJAKARA UKARAM = Tamil letter nju
ங்	9B	TAMIL LETTER NGAKARA UUKAARAM = Tamil letter nguu
ஞ்	9C	TAMIL LETTER NJAKARA UUKAARAM = Tamil letter njuu

௩	9D	TAMIL NUMBER PATHTHU = Tamil number ten
௩	9E	TAMIL NUMBER NUURRU = Tamil number one hundred
௩	9F	TAMIL NUMBER AAYIRAM = Tamil number one thousand
	A0	<NOT ASSIGNED>
ஶ	A1	TAMIL VOWEL SIGN KAAL = Tamil vowel sign aa
ஶ	A2	TAMIL VOWEL SIGN KOKKI = Tamil vowel sign i
ஶ	A3	TAMIL VOWEL SIGN CUZHI-K-KOKKI = Tamil vowel sign ii
ஶ	A4	TAMIL VOWEL SIGN MEEL VILANGU or KONNNTAI = Tamil vowel sign u
ஶ	A5	TAMIL VOWEL SIGN CUZHI MEEL VILANGU or CUZHIK KONNNTAI = Tamil vowel sign uu
ஶ	A6	TAMIL VOWEL SIGN KOMPU = Tamil vowel sign e
ஶ	A7	TAMIL VOWEL SIGN IRATTAI-K-KOMPU = Tamil vowel sign ee
ஶ	A8	TAMIL VOWEL SIGN IRATTAI-C-CUZHI = Tamil vowel sign ai
©	A9	COPYRIGHT SIGN
ஶ	AA	TAMIL VOWEL SIGN CIRRAKU = Tamil au length mark
அ	AB	TAMIL LETTER AKARAM = Tamil letter a
ஆ	AC	TAMIL LETTER AAKAARAM = Tamil letter aa

இ	AD	TAMIL VOWEL IKARAM (USAGE IN SLOT DEPRECATED) = Tamil letter i (usage in slot deprecated)
ஈ	AE	TAMIL LETTER IIKAARAM = Tamil letter ii
உ	AF	TAMIL LETTER UKARAM = Tamil letter u
ஊ	B0	TAMIL LETTER UUKAARAM = Tamil letter uu
எ	B1	TAMIL LETTER EKARAM = Tamil letter e
ஏ	B2	TAMIL LETTER EEKAARAM = Tamil letter ee
ஐ	B3	TAMIL LETTER AIKAARAM = Tamil letter ai
ஓ	B4	TAMIL LETTER OKARAM = Tamil letter o
ஔ	B5	TAMIL LETTER OOKAARAM = Tamil letter oo
ஔள	B6	TAMIL LETTER AUKAARAM = Tamil letter au
ஃ	B7	TAMIL AAYTHAM LETTER AKHENAM or AKHAAN = Tamil letter aaytham
க	B8	TAMIL LETTER KAKARA AKARAM = Tamil letter ka
ங	B9	TAMIL LETTER NGAKARA AKARAM = Tamil letter nga
ச	BA	TAMIL LETTER CAKARA AKARAM = Tamil letter ca
ஞ	BB	TAMIL LETTER NJAKARA AKARAM = Tamil letter nja

L	BC	TAMIL LETTER TAKARA AKARAM = Tamil letter tta
ண	BD	TAMIL LETTER NNNAKARA AKARAM = Tamil letter nna
த	BE	TAMIL LETTER THAKARA AKARAM = Tamil letter ta
ந	BF	TAMIL LETTER NAKARA AKARAM = Tamil letter na
ப	C0	TAMIL LETTER PAKARA AKARAM = Tamil letter pa
ம	C1	TAMIL LETTER MAKARA AKARAM = Tamil letter ma
ய	C2	TAMIL LETTER YAKARA AKARAM = Tamil letter ya
ர	C3	TAMIL LETTER RAKARA AKARAM = Tamil letter ra
ல	C4	TAMIL LETTER LAKARA AKARAM = Tamil letter la
வ	C5	TAMIL LETTER VAKARA AKARAM = Tamil letter va
ழ	C6	TAMIL LETTER ZHAKARA AKARAM = Tamil letter lla
ள	C7	TAMIL LETTER LLAKARA AKARAM = Tamil letter lla
ற	C8	TAMIL LETTER RRAKARA AKARAM = Tamil letter rra
ன	C9	TAMIL LETTER NNAKARA AKARAM = Tamil letter nna
டி	CA	TAMIL LETTER TAKARA IKARAM = Tamil letter tti

ஐ	CB	TAMIL LETTER TAKARA IIKAARAM = Tamil letter tti
ஓ	CC	TAMIL LETTER KAKARA UKARAM = Tamil letter ku
ஔ	CD	TAMIL LETTER CAKARA UKARAM = Tamil letter cu
ஐ	CE	TAMIL LETTER TAKARA UKARAM = Tamil letter ttu
ஊ	CF	TAMIL LETTER NNNAKARA UKARAM = Tamil letter nnnu
து	D0	TAMIL LETTER THAKARA UKARAM = Tamil letter tu
து	D1	TAMIL LETTER NAKARA UKARAM = Tamil letter nu
பு	D2	TAMIL LETTER PAKARA UKARAM = Tamil letter pu
மு	D3	TAMIL LETTER MAKARA UKARAM = Tamil letter mu
யு	D4	TAMIL LETTER YAKARA UKARAM = Tamil letter yu
ரு	D5	TAMIL LETTER RAKARA UKARAM = Tamil letter ru
லு	D6	TAMIL LETTER LAKARA UKARAM = Tamil letter lu
வு	D7	TAMIL LETTER VAKARA UKARAM = Tamil letter vu
ழு	D8	TAMIL LETTER ZHAKARA UKARAM = Tamil letter llu
ளு	D9	TAMIL LETTER LLAKARA UKARAM = Tamil letter llu

று	DA	TAMIL LETTER RRAKARA UKARAM = Tamil letter rru
னு	DB	TAMIL LETTER NNAKARA UKARAM = Tamil letter nnu
கூ	DC	TAMIL LETTER KAKARA UUKAARAM = Tamil letter kuu
சூ	DD	TAMIL LETTER CAKARA UUKAARAM = Tamil letter cuu
஠	DE	TAMIL LETTER TAKARA UUKAARAM = Tamil letter ttuu
ணூ	DF	TAMIL LETTER NNNAKARA UUKAARAM = Tamil letter nnnuu
தூ	E0	TAMIL LETTER THAKARA UUKAARAM = Tamil letter tuu
நூ	E1	TAMIL LETTER NAKARA UUKAARAM = Tamil letter nuu
பூ	E2	TAMIL LETTER PAKARA UUKAARAM = Tamil letter puu
மூ	E3	TAMIL LETTER MAKARA UUKAARAM = Tamil letter muu
யூ	E4	TAMIL LETTER YAKARA UUKAARAM = Tamil letter yuu
ரூ	E5	TAMIL LETTER RAKARA UUKAARAM = Tamil letter ruu
லூ	E6	TAMIL LETTER LAKARA UUKAARAM = Tamil letter luu
வூ	E7	TAMIL LETTER VAKARA UUKAARAM = Tamil letter vuu
ழூ	E8	TAMIL LETTER ZHAKARA UUKAARAM = Tamil letter lluu



஌	E9	TAMIL LETTER LAKARA UUKAARAM = Tamil letter lluu
றா	EA	TAMIL LETTER RRAKARA UUKAARAM = Tamil letter rruu
னா	EB	TAMIL LETTER NNAKARA UUKAARAM = Tamil letter nnuu
க்	EC	TAMIL LETTER KAKARAM = Tamil letter k
ங்	ED	TAMIL LETTER NGAKARAM = Tamil letter ng
ச்	EE	TAMIL LETTER CAKARAM = Tamil letter c
ஞ்	EF	TAMIL LETTER NJAKARAM = Tamil letter nj
ட்	F0	TAMIL LETTER TAKARAM = Tamil letter tt
ண்	F1	TAMIL LETTER NNNAKARAM = Tamil letter nnn
த்	F2	TAMIL LETTER THAKARAM = Tamil letter t
ந்	F3	TAMIL LETTER NAKARAM = Tamil letter n
ப்	F4	TAMIL LETTER PAKARAM = Tamil letter p
ம்	F5	TAMIL LETTER MAKARAM = Tamil letter m
ய்	F6	TAMIL LETTER YAKARAM = Tamil letter y
ர்	F7	TAMIL LETTER RAKARAM = Tamil letter r

ல்	F8	TAMIL LETTER LAKARAM = Tamil letter l
வ்	F9	TAMIL LETTER VAKARAM = Tamil letter v
ழ்	FA	TAMIL LETTER ZHAKARAM = Tamil letter LLL
ள்	FB	TAMIL LETTER LLAkARAM = Tamil letter ll
ற்	FC	TAMIL LETTER RRAkARAM = Tamil letter rr
ன்	FD	TAMIL LETTER NNAkARAM = Tamil letter nn
இ	FE	TAMIL LETTER IKARAM = Tamil letter i
	FF	<NOT ASSIGNED>

## NOTES:

Third vowel "i" is included in slots AD and EF but usage of "i" at slot AD is deprecated. Inclusion of the glyph at slot AD is for rendering legacy data and to enable conversion to other encodings. Text converters to other encodings should attempt to determine which slot is used in the text for ikaram before converting.

Though ukara- and ukara modifiers (at slots 4A, 4B) are indicated as "TAMIL VOWEL SIGN U" and "TAMIL VOWEL SIGN UU" respectively, their usage is permitted only for the grantha vowels. Entire ukara and ukara uyirmey series are encoded directly in TSCII and they alone are to be used to render these uyirmeys.

Tamil numerals 0-9 are indicated as "TAMIL DIGITS" while Tamil numerals 10,100 and 1000 are indicated as "TAMIL NUMBERS". This is to recognize the fact that Tamil numerals are being used in two different systems (decimalic as in Arabic using digits 0-9 and as an additive-positional system using numerals 10,100 and 1000 as well).

TAMIL LETTER AAYTHAM at slot 7B is a dependant letter though in Unicode 4.1 it is listed differently as TAMIL VISARGA SIGN. (in Tamil grammar this aaytham letter is called as "caarpu ezuttu"),