

GRETIL Concordance of Encodings / Transliteration Systems

		CSX+		REE		UTF-8	ATIL	Aiba	KyH
´	<i>acute (for udātta)</i>						´		
'	<i>single quote (for avagraha)</i>	039	'	039	'	᳚	'		^
–	<i>hyphen</i>	045	–	045	–	ΓÇÉ	–		
`	<i>grave (for svarita)</i>	096	`	096	`	᳚	`		
Ç	<i>C cedilla</i>	128	Ç	128	Ç	┌ç	,C		
ü	<i>u dieresis</i>	129	ü	129	ü	┌ll	:u		
é	<i>e acute</i>	130	é	130	é	┌r	!e		
â	<i>a circumflex</i>	131	â	131	â	┌ó	^a		
ä	<i>a dieresis</i>	132	ä	132	ä	┌ñ	:a		
à	<i>a grave</i>	133	à	133	à	┌á	`a		
å	<i>a ring</i>	134	å	134	å	┌Ñ	@a		
ç	<i>c cedille</i>	135	ç	135	ç	┌°	,c		
ê	<i>e circumflex</i>	136	ê	136	ê	┌r	^e		
ë	<i>e dieresis</i>	137	ë	137	ë	┌½	:e		
è	<i>e grave</i>	138	è	138	è	┌z	`e		
ï	<i>i dieresis</i>	139	ï	139	ï	┌»	:i		
î	<i>i circumflex</i>	140	î	140	î	┌«	^i		
ì	<i>i grave</i>	141	ì	141	ì	┌¼	`i		
Ä	<i>A dieresis</i>	142	Ä	142	Ä	┌ä	:A		
Å	<i>A ring</i>	143	Å	143	Å	┌à	@A		
É	<i>E acute</i>	144	É	144	É	┌ë	!E		
æ	<i>ae</i>	145	æ	145	æ	┌ ^a	ae		
Æ	<i>AE</i>	146	Æ	146	Æ	┌å	AE		
ô	<i>o circumflex</i>	147	ô	147	ô	┌	^o		
ö	<i>o dieresis</i>	148	ö	148	ö	┌ll	:o		
ò	<i>o grave</i>	149	ò	149	ò	┌	`o		
û	<i>u circumflex</i>	150	û	150	û	┌r	^u		
ù	<i>u grave</i>	151	ù	151	ù	┌ll	`u		
ā	<i>ae macron</i>	152	ij			┌ú	aeae		

		CSX+		REE		UTF-8	ATIL	Aiba	KyH
Ö	<i>O dieresis</i>	153	Ö	153	Ö	┌û	:O		
Ü	<i>U dieresis</i>	154	Ü	154	Ü	┌œ	:U		
Û	<i>u breve</i>	155	ϣ			┌i	;u		
Ě	<i>e macron tilde</i>	156	£			—ô Ꞥâ	~ee		
ŕ	<i>r underring *</i>	157	¥			r ꞤÑ	:r		
á	<i>a acute</i>	158	Pt			┌í	!a		
ṙ	<i>r underbar</i>	159	f			βꞤ f	_r		
í	<i>i acute</i>	161	í	161	í	┌i	!i		
ó	<i>o acute</i>	162	ó	162	ó	┌	!o		
ú	<i>u acute</i>	163	ú	163	ú	┌	!u		
ñ	<i>n tilde</i>	164	ñ	164	ñ	┌⋮	~n	~n	J
Ñ	<i>N tilde</i>	165	Ñ	165	Ñ	┌æ	~N		
İ	<i>l tilde</i>	166	ª			l Ꞥâ	~l		
ṁ	<i>m overdot *</i>	167	°			βꞤ ü	^m		
ǎ	<i>a macron breve</i>	168	¿			—ü Ꞥâ	;aa		
ĩ	<i>i macron breve</i>	169	¬			—½ Ꞥâ	;ii		
ũ	<i>u macron breve</i>	170	¬			┌½ Ꞥâ	;uu		
ã	<i>a macron tilde</i>	171	½			—ü Ꞥâ	~aa		
ĩ	<i>i macron tilde</i>	172	¼			—½ Ꞥâ	~ii		
ṅ	<i>n underbar</i>	173	i			βꞤ ë	_n		
Ṛ	<i>r underring macron *</i>	174	«			r ꞤÑ Ꞥä	:rr		
ḷ	<i>l underring *</i>	175	»			l ꞤÑ	:l		
Ḹ	<i>l underring macron *</i>	176	⋮			l ꞤÑ Ꞥä	:ll		
ŕ	<i>r underring acute *</i>	177	⋮			r ꞤÑ Ꞥü	!r		
ṙ	<i>r underring grave *</i>	178	⋮			r ꞤÑ Ꞥç	`r		
Ṛ	<i>r underring macron acute *</i>	179				r ꞤÑ Ꞥä Ꞥü	!rr		
ṁ	<i>m breve</i>	180	┌			m Ꞥâ	;m		
á	<i>a macron acute</i>	181	≡			—ü Ꞥü	!aa		
à	<i>a macron grave</i>	182	≡			—ü Ꞥç	`aa		

		CSX+		REE	UTF-8	ATIL	Aiba	KyH
í	<i>i macron acute</i>	183	⌈		—½ ⌈ü	lii		
ì	<i>i macron grave</i>	184	⌋		—½ ⌋Ç	`ii		
ē	<i>e macron</i>	185	⌈		—ô	ee		
ō	<i>o macron</i>	186	⌈		+î	oo		
Ŕ	<i>R undererring *</i>	187	⌋		R ⌋Ñ	:R		
ÿ	<i>y overdot</i>	188	⌋		β ⌈ Å	.y		
ú	<i>u macron acute</i>	189	⌋		+½ ⌋ü	!uu		
ù	<i>u macron grave</i>	190	⌋		+½ ⌋Ç	`uu		
ř	<i>r breve</i>	191	⌋		r ⌋â	;r		
ō	<i>o macron tilde</i>	192	⌋		+î ⌋â	~oo		
ṁ	<i>m candrabindu *</i>	193	⌋		m ⌋É / m ⌋Ç	~m		
ṭ	<i>t underbar</i>	194	⌋		β⌋ »	_t		
Ē	<i>E macron</i>	195	⌋		—Æ	EE		
Ō	<i>O macron</i>	196	—		+î	OO		
ň	<i>n breve</i>	197	+		n ⌋â	;n		
ř	<i>r underdot acute</i>	198	⌋		β⌋ ç ⌋ü	!r		
ṛ	<i>r underdot grave</i>	199	⌋		β⌋ ç ⌋Ç	`r		
<u>Kh</u>	<i>Kh underbar</i>	200	⌋		KΓÇih ⌋	_Kh		
<u>k</u>	<i>k underbar</i>	201	⌋		β⌋ ⌋	_k		
ĀE	<i>AE macron</i>	203	⌋		⌋ó	AEAE		
<u>kh</u>	<i>kh underbar</i>	204	⌋		kΓÇih ⌋	_kh		
ġ	<i>g overdot</i>	205	=		—í	.g		
ĉ	<i>c circumflex</i>	206	⌋		—ë	^c		
ř	<i>r underdot macron acute</i>	207	⌋		β⌋ ¥ ⌋ü	!..rr		
ã	<i>a tilde</i>	208	⌋		⌋ú	~a		
ĩ	<i>i tilde</i>	209	⌋		—r	~i		
ũ	<i>u tilde</i>	210	⌋		+r	~u		
ẽ	<i>e tilde</i>	211	⌋		β ⌋ ⌋	~e		
õ	<i>o tilde</i>	212	⌋		⌋	~o		

		CSX+		REE		UTF-8	ATIL	Aiba	KyH
ě	<i>e breve</i>	213	Ɔ			—ò	;e		
ö	<i>o breve</i>	214	Ɔ			—Å	;o		
ḷ	<i>l underbar</i>	215	Ɔ			β̣ ̣	_l		
ũ	<i>u macron tilde</i>	216	≠			—½ ̣	~uu		
Ġ	<i>G overdot</i>	217	┘			—á	.G		
Ĉ	<i>C circumflex</i>	218	┘			—ê	^C		
ḥ	<i>h underbar</i>	219	■			β û	_h		
ḥ	<i>h underbreve</i>	220	■			β̣ ½	;h		
–	<i>endash</i>	221	■			ΓÇô	[-]		
—	<i>emdash</i>	222	■			ΓÇö	[- -]		
“	<i>quote double left</i>	223	■			ΓÇ£	“*		
ā	<i>a macron</i>	224	α	195	┘	—ü	aa	aa	A/aa
ß	<i>German double s</i>	225	β	225	β	┘f	ss		
Ā	<i>A macron</i>	226	Γ	249	·	—Ç	AA		
ī	<i>i macron</i>	227	π	197	+	—½	ii	ii	I/ii
Ī	<i>I macron</i>	228	Σ	253	²	—┘	II		
ū	<i>u macron</i>	229	σ	198	≠	—½	uu	uu	U/uu
Ū	<i>U macron</i>	230	μ	224	α	—┘	UU		
ṛ	<i>r underdot</i>	231	τ	173	i	β̣ ̣	.r	.r	R
Ṛ	<i>R underdot</i>	232	Φ	227	π	β̣ ̣ Ü	.R		
ṝ	<i>r underdot macron</i>	233	■	204	≠	β̣ ̣ ¥	..rr	~r	RR/Q
Ṝ	<i>R underdot macron</i>	234	Ω			β̣ ̣ £	..RR		
ḷ	<i>l underdot</i>	235	δ	202	≠	β̣ ̣ ̣	.l	.l	L/R
Ḷ	<i>L underdot</i>	236	∞			β̣ ̣ ̣	.L		
ḹ	<i>l underdot macron</i>	237	φ	203	≠	β̣ ̣ ̣	..ll	~l	LL/IRR
Ḹ	<i>L underdot macron</i>	238	ε			β̣ ̣ ̣	..LL		
ṅ	<i>n overdot</i>	239	∩	199	≠	β̣ ̣ à	^n	^n	G
Ñ	<i>N overdot</i>	240	≡	167	°	β̣ ̣ ä	^N		
ṭ	<i>t underdot</i>	241	±	194	┘	β̣ ̣ i	.t	.t	T

		CSX+		REE		UTF-8	ATIL	Aiba	KyH
Ṭ	<i>T underdot</i>	242	≥	232	Φ	β≡ ¼	.T		
ḍ	<i>d underdot</i>	243	≤	172	¼	β≡ ì	.d	.d	D
Ḍ	<i>D underdot</i>	244	∫	214	∏	β≡ î	.D		
ṇ	<i>n underdot</i>	245	J	239	∩	β≡ ç	.n	.n	N
Ṇ	<i>N underdot</i>	246	÷	215	≡	β≡ â	.N		
ś	<i>s acute</i>	247	≈	211	⊥	†ç	^s	^s	z/Z
Ś	<i>S acute</i>	248	°	193	⊥	†Ü	^S		
ṣ	<i>s underdot</i>	249	·	171	½	β≡ ú	.s	.s	S
Ṣ	<i>S underdot</i>	250	·	229	σ	β≡ ó	.S		
”	<i>quote double right</i>	251	√			ΓÇ¥	*"		
ṡ	<i>m underdot</i>	252	ⁿ	230	μ	β≡ â	.m	.m	M
Ṣ	<i>M underdot</i>	253	²	245	J	β≡ é	.M		
ḥ	<i>h underdot</i>	254	■	247	≈	β≡ Ñ	.h	.h	H
Ḥ	<i>H underdot</i>	255					.H		
h	<i>h in aspirates (optional)</i>					≡			

- ◇ The list follows the CSX+ (Classical Sanskrit eXtended+) encoding table (in ASCII sequence). Because CSX+ covers most aspects indologists may need for the practical purpose of entering and handling e-texts in Indian languages, it remains a valuable reference point.
Character descriptions according to John Smith’s list of CSX+ encoding:
<ftp://bombay.oriental.cam.ac.uk/pub/john/software/fonts/csx+/CSX+.def>
- ◇ Grey shading indicates the GRETIL standard character set.
- ◇ Asterisked characters (*) are replaced with the respective underdot variants in the GRETIL standard character set.
- ◇ Blank fields indicate that the characters are not defined in the respective encoding / transliteration system.
- ◇ CSX+, REE and UTF-8 characters are displayed according to codepage 437; the UTF-8 column gives the respective multibyte sequences.
- ◇ ATIL, the ASCII Transliteration for Indian Languages devised for GRETIL on the basis of pure ASCII code (ISO 646), takes up elements of the present transliteration practice of Muneo Tokunaga, Masato Fujii, Tooru Aiba and other Japanese scholars, to whose unselfish efforts the indological community owes so many e-texts.
The distinctive feature of ATIL is that it covers the entire range of characters defined in CSX+ encoding, even though some of them may seldom be used by indologists.

- ◇ For a comprehensive chart of Aiba and Kyoto-Harvard (KyH) transliterations see <http://texa.human.is.tohoku.ac.jp/aiba/codes/table-skt-01.html>

Not yet defined in ATIL:

- ◇ Discrimination of vedic hiatus / Prakrit juxtaposition of vowels
- ◇ Initial vowels in Tamil

For a systematic list of GRETIL diacritics see http://www.sub.uni-goettingen.de/ebene_1/findolo/gretil/gretdias.pdf.

Reinhold Grünendahl, 6.12.2002
(gruenen@sub.uni-goettingen.de)