

Malayalam Chillus in grapheme clusters

Cibu Johny, Umesh Nair
 cibu@google.com, umesh@google.com
 2007-10-19

The table below displays various NA/Chillu-N, RRA combinations. Alpha, Beta and Gamma columns are plausible storage representations.

Beta can conveniently represent all forms; however, distinguishes row A and row B which are same in reading. It considers Chillu as a consonant graphically and makes use of <chillu, virama> sequence. This however, does not follow user expected order for encoding, in case of row B.

Alpha does not use <chillu, virama> sequence and distinguishes the rows with same reading (A and B) by only ZWJ. However, it does not have a reasonable representation for row E.

Gamma column is author's capture from UTC meeting in Feb 2007.

Table 1	Visuals	α	β	γ	Reading
A	നേ	ന, ്, റ, െ (NA, VIRAMA, RRA, E) (0D28, 0D4D, 0D31, 0D46)	നേ, ്, റ, െ (Chillu-N, VIRAMA, RRA, E) (0D7B, 0D4D, 0D31, 0D46)	same as β	nte
B	നേറ	ന, ്, ZWJ, റ, െ (NA, VIRAMA, ZWJ, RRA, E) (0D28, 0D4D, 200D, 0D31, 0D46)	നേ, െ, റ, (Chillu-N, E, RRA) (0D7B, 0D46, 0D31)	നേ, ്, ZWJ, റ, െ (Chillu-N, VIRAMA, ZWJ, RRA, E) (0D7B, 0D4D, 200D, 0D31, 0D46)	nte
C	നേറ	നേ, റ, െ (Chillu-N, RRA, E) (0D7B, 0D31, 0D46)	same as α	same as α	nrre
E	നേററ	?	നേ, െ, റ, റ, (Chillu-N, E, RRA, RRA) (0D7B, 0D46, 0D31, 0D31)	?	nte
F	നേറ	ന, ്, റ, െ (NA, VIRAMA, RRA, E) (0D28, 0D4D, 0D31, 0D46)	same as α	same as α	nərre
G	നേറ	ന, റ, െ (NA, RRA, E) (0D28, 0D31, 0D46)	same as α	same as α	narre

Following table lists various RRA and RRA combinations. All the cases are quite similar to Table 1 where NA, RRA combinations are discussed. Rows X, Y and Z match rows A, B and C respectively. Basic principles and properties of Alpha and Beta columns remain the same.

Table 2	Visuals	β		
		α	(current representation)	Reading
X		o, ʼ, o, ʼ (RRA, VIRAMA, RRA, E) (0D31, 0D4D, 0D31, 0D46)	same as α	te
Y		o, ʼ, ZWJ, o, ʼ (RRA, VIRAMA, ZWJ, RRA, E) (0D31, 0D4D, 200D, 0D31, 0D46)	o, ʼ, o (RRA, E, RRA) (0D31, 0D46, 0D31)	'te' or 'rrerra'. 'rrerra' is unusual reading.
Z		o, o, ʼ (RRA, RRA, E) (0D31, 0D31, 0D46)	same as α	rrarre

The conjuncts with dot-reph and corresponding ones with chillu-R are *almost* always read the same. So user expectation is that both are interchangeable even though there are few exceptions.

The Alpha column uses the same model as rest of the indic: <RA, virama, consonant>. Alpha column, however, needs ZWJ to represent the very frequent conjunct <RA, virama, YA>. Beta column uses the <chillu, virama> paradigm concurring with user expectation.

Table 3	#	Visual	α	β1	reading	notes
ᱠ(RA - 0D30) with ᱡ(YA - 0D2F)	Y1	ᱠᱟ	ᱠ, ᱟ, ᱡ (RA, Virama, YA) (0D30, 0D4D, 0D2F)	ᱠᱟ, ᱟ, ᱡ (Chillu-R, Virama, YA) (0D7C, 0D4D, 0D2F)	rrya	No examples found yet.
	Y2	ᱠᱟᱡ	ᱠᱟ, ᱡ (Chillu-R, YA) (0D7C, 0D2F)	same as α	rrya	
	Y3	ᱠᱟᱪ	ᱠ, ᱟ, ZWJ, ᱡ (RA, Virama, ZWJ, YA) (0D30, 0D4D, 200D, 0D2F)	ᱠ, ᱟ, ᱡ (RA, Virama, YA) (0D30, 0D4D, 0D2F)	rya	very frequent
	Y4	ᱠᱟᱪᱟ	ᱠ, ᱟ, ZWNJ, ᱡ (RA, Virama, ZWNJ, YA) (0D30, 0D4D, 200C, 0D2F)	same as α	rəya	No examples found yet.
	Y5	ᱠᱟᱪᱟᱪ	ᱠ, ᱟ, ᱡ, ᱟ, ᱡ (RA, Virama, YA, Virama, YA) (0D30, 0D4D, 0D2F, 0D4D, 0D2F)	ᱠᱟ, ᱟ, ᱡ, ᱟ, ᱡ (Chillu-R, Virama, YA, Virama, YA) (0D7C, 0D4D, 0D2F, 0D4D, 0D2F)	rryya	
	Y6	ᱠᱟᱪᱟᱪᱟ	ᱠᱟ, ᱡ, ᱟ, ᱡ (Chillu-R, YA, Virama, YA) (0D7C, 0D2F, 0D4D, 0D2F)	same as α	rryya	No examples found yet.
ᱠ(RA - 0D30) with ᱢ(VA - 0D35)	V1	ᱠᱟᱢ	ᱠ, ᱟ, ᱢ (RA, Virama, VA) (0D30, 0D4D, 0D35)	ᱠᱟ, ᱟ, ᱢ (chillu-R, Virama, VA) (0D7C, 0D4D, 0D35)	rrva	
	V2	ᱠᱟᱢᱟ	ᱠᱟ, ᱢ (Chillu-RA, VA) (0D7C, 0D35)	same as α	rrva	
	V3	ᱠᱟᱢᱟᱪ	ᱠ, ᱟ, ZWJ, ᱢ (RA, Virama, ZWJ, VA) (0D30, 0D4D, 200D, 0D35)	ᱠ, ᱟ, ᱢ (RA, Virama, VA) (0D30, 0D4D, 0D35)	rva	To represent colloquial tongue

	V4	രവ	ര, ്, ZWNJ, വ (RA, Virama, ZWNJ, VA) (0D30, 0D4D, 200C, 0D35)	same as α	rava	No examples found yet.
	V5	വ്വ	ര, ്, വ, ്, വ (RA, Virama, VA, Virama, VA) (0D30, 0D4D, 0D35, 0D4D, 0D35)	ര, ്, വ, ്, വ (Chillu-R, Virama, VA, Virama, VA) (0D7C, 0D4D, 0D35, 0D4D, 0D35)	rrvva	
	V6	രവ്വ	ര, വ, ്, വ (Chillu-R, VA, Virama, VA) (0D7C, 0D35, 0D4D, 0D35)	same as α	rrvva	
RA with other consonants	S1	ഷ	ര, ്, ഷ (RA, Virama, SSA) (0D30, 0D4D, 0D37)	ര, ്, ഷ (Chillu-R, Virama, SSA) (0D7C, 0D4D, 0D2F)	rrsha	

Chillus embedded grapheme clusters are used in [A Grammar of the Malayalam Language](#) by Hermann Gundert (year 1868).

42. ടകാരം സവണ്ണങ്ങൾക്കും ഷകാരത്തിന്നും പകരം (ശബ്ദം-ചഞ്ച; ഡക-ഇടക്ക; ഖണ്ഡം-കഞ്ചം; മേഘം-മേടം; പൂർവ്വം-പൂർവ്വം; ഗോഷം-ഗോട്ടം) ഷഡ്ഗാഹം രാട്ട മുതലായവറ്റിൽ ഉകാരം അധികം നടപ്പു. (കൾ-രാൾ). പിന്നെ മലയാള ടകാരം പലതും നെളകാരങ്ങളിൽ നിന്നുജനിച്ചവ (ഇരട്ട-രട്ട്; കാട്ടുക-ടുക; കേട്ട-കേറ്റ)

Conjunct highlighted above has Chillu-LL and <TA, VIRAMA, TA, U-Sign> ligature.

1.) ഉ-ം കൂതത്തോടൊപരകലി എന്നിങ്ങനെ ര യുഗന്തവികലും (കേ. ഉ.) ഗുരുനം അതിസാരവും വിചം കൈവിചം സവ്വ രോഗത്തിന്നും നന്ന (വൈ. ശം.)
 2.) പട്ടത്തിൽ മുന്നേത്ത നാമങ്ങളും വിഭക്തിക്കുകയാണു. ഉ-ം പക്ഷിപ്പക്ഷികൾ എന്നു കേയും പറയുമ്പോൾ ആട്ടിന്നും മൂലാദികൾ എന്നിവ പതിനാറു വയസ്സ് (വൈ. ച). 395. കാക്ക.

(Page 107)

Conjunct highlighted above has Chillu-LL and <KA, VIRAMA, KA> ligature.