

Proposal to encode the Four Diacritic Marks of Kannada in Unicode

Biswajit Mandal (biswajitmandal.bm90@gmail.com)

Mar 28, 2026

This is a proposal to encode four new characters in the Kannada block of the Unicode standard:

GLYPH	CODE	CHARACTER NAME
◌̄	U+0CD0	KANNADA SIGN OVERLINE
◌̇	U+0CD1	KANNADA SIGN VERTICAL LINE ABOVE
◌̈	U+0CD2	KANNADA SIGN DOT ABOVE
◌̆	U+0CD3	KANNADA SIGN CANDRA ABOVE

The Kodava (ಕೊಡವ ತಕ್ಕು [koḍəvə t̪ək:i]) ISO 639-3 kfa, meaning 'speech of Kodavas' is a Dravidian language spoken in Kodagu district (Coorg) in Southern Karnataka, India. It is an endangered language. The 2011 Census of India reports 96,918 persons speak Kodava as their mother tongue. Kodava is currently written in Kannada script with some extra diacritical marks for unrounded phonemes which are not present in Kannada script. Now, some new scripts were invented and one of them selected for official script, but the progress is in initial stage. In some books these diacritical marks were used as to show unrounded phonemes that is [i], [ɨ], [ɜ] and [ɛ]. But the position and usage were not clear in those books. After NEP 2020 (New Education Policy) of India; it emphasised on the mother tongue education in the foundational and preparatory stages (elementary stages) of the child. On the basis of this policy NCERT (National Council of Educational Research and Training) developed a primer for Kodava clearly mentioned the usages and position of these diacritics.



These marks also occur with the vowel signs:

ಕೊಡವ ವರ್ಣಮಾಲೆ

ಪ್ರಾರಂಭ

ಅ	ಆ	ಇ	ಉ	ಊ	ಋ	
ಋ	ೠ	ಎ	ಏ	ಒ	ಓ	
ಓ	ಔ	ಐ	ಒ	ಓ	ಔ	ಔ
ಔ						

Fig 1: Kodava Vowels. (NCERT & CIIL 2024:1)

	<p>ಅಗೆ</p>	
<p>ನಾಟಕ್ ಬೋಂಡು ಅಗೆ ಅಣ್ಣಯ್ಯ ಬೆಚ್ಚತ್ ಪಗೆ ಬೇಲ್‌ಕ್ ಎಡ್‌ಕಂಡು ಬಗೆ</p>		
		

Fig 2: usage of dot above and vertical line. (NCERT & CIIL 2024: 3)

<p style="text-align: center;">— ಇ</p>	<p style="text-align: center;">ಇಟ್ಟಿಗೆ</p>				
<p>ಮನೆಕಟ್ಟೋಕ್ ಬೋಂಡು ಇಟ್ಟಿಗೆ ಮಂಡೆಕ್ ಚೂಡೋ ಮಲ್ಲಿಗೆ ತಿಂಬಕ್ ಚೋದ್ ನೆಲ್ಲಿಗೆ</p>					
					
<p>ಇರಾಟ್</p>	<p>ಕತ್ತಲೆ</p>	<p>ಇಳಿ</p>	<p>ಇಳಿ</p>	<p>ಇಡ್</p>	<p>ತೊಡು</p>

Fig 3: Usage of Overline, dot above and vertical bar. (NCERT & CIIL 2024:07)

		<h1>ಚೀಲೆ</h1>			
		<h1>ದಟ್ಟ</h1>			
<p>ಕುಪ್ಪಿಯಕ್ ಕೆಟ್ಟಂಡು ಚೀಲೆ ಕೂಳಿತೆರೆಕಾಂಡು ಪಾಲೆ ಪೊಂಗಡ ಕೆಮಿಕಾಂಡು ವಾಲೆ</p>					
					
ತೇಲೆಮರ	ಬಸರಿಮರ	ತೇರಣ	ಮದರಂಗಿ	ಬೇಳೆ	ಬೇಳೆ

Fig 4: usage diacritics with vowel signs. (NCERT & CIIL 2024:16)

	<p>ಋಷೋಃ</p>				
<p>ಕಾಪಳತೊರ್ ಪೋಡೆ ಚುಣ್ಣ ಪೊಜ್ಜೊರ್ ಗೋಡೆ ಸೇರ್ಲ್ ಅಳಕಂಡು ಮೋಡೆ</p>					
					
<p>ಋಷೋಃ</p>	<p>ಕುತ್ತಿಗೆ</p>	<p>ಮೋಡ</p>	<p>ಮೋಡ</p>	<p>ಪೋಡೆ</p>	<p>ಹೇಂಟೆ</p>

Fig 5: Use of candra above with letters and vowel signs too. (NCERT & CIIL 2024:21)



<p style="text-align: center;">। ಎ</p>	<p style="text-align: center;">ಎಟ್ತ್ಯೆ</p>				
<p>ನಾಲ್ ನಾಲ್ ಕೂಡಿ ಎಟ್ತ್ಯೆ ಬೋಟೆಲ್ ಎಡ್ ನೀ ಬಟ್ತ್ಯೆ ಉಂಡಿತ್ ಪತ್ತ್ ನೀ ಕಟ್ತ್ಯೆ</p>	<p style="text-align: center;">8</p>				
					
<p>ಎಡ್</p>	<p>ತೆಗಿ / ತೆಗೆದುಕೋ</p>	<p>ಎಳ್ತಾ</p>	<p>ಬರವಣಿಗೆ</p>	<p>ಎಡೆ</p>	<p>ಎಡೆ</p>

Fig 6: Use of vertical bar. (NCERT & CIIL 2024:13)

ತಟ್ಟುಮುಟ್ಟು; 'ಎಡೆತಡೆ; ತೊಂದರೆ.	ಅಡಿ ² (ನಾ): one foot; a measurement of twelve inches: ಅಡಿ; ಪನ್ನೆರಂಡ್ ಇಂಚ್ ಅಳ್ತೆ.
ಅಡಪ (ನಾ): small bag or pouch for keeping arecanuts and betel leaves: ಉರುಪೆ; ಕೊಡಿಯಲಡಿಕೆ ಚೀಲ.	ಅಡಿ ³ (ನಾ): bottom; base: ಅಡಿ; ಕೇದ ಬರಿ; ತವ್ವತ ಬರಿ.
ಅಡಮಾನ (ನಾ): pledge; mortgage; pawn: ಅಡ; ಅಡವ್; ಗಿರ್ವಿ.	ಅಡಿಕೆ = ಅಡಕೆ.
ಅಡರು ¹ (ಕ್ರಿ): mount; ascend; climb: ಕೊಡಿಕ ಪತ್ತ; ಕೊಡಿಕ ಪೋ.	ಅಡಿಕೆಮರ = ಅಡಕೆಮರ.
ಅಡರು ² (ಕ್ರಿ): spread; pervade through: ಪಬ್ಬ; ಪರ್ಡ್.	ಅಡಿಗಟ್ಟು (ನಾ): foundation; base: ಪಾಯ; ಅಡಿಪಾಯ; ಮನೆತರೆ.
ಅಡರುಬಳ್ಳಿ (ನಾ): climber: ತಡೆ; ಬಣ್ಣತ ಬಳ್ಳಿ.	ಅಡಿಗಲ್ಲು (ನಾ): foundation stone: ಪಾಯಕಲ್ಲ; ಅಡಿಪಾಯತರ ಕಲ್ಲ.
ಅಡವಿ (ನಾ): forest; jungle: ಕಾಡ್; ಬನ.	ಅಡಿಗಾಲುವೆ (ನಾ): culvert: ಮೋರಿ; ರಸ್ತೆರ ಅಡಿಲ್ ನೀರ್ ಪರಿವಕ್ ಅಡ್ಡ ಕೊಳಪೆ ಬೆಚ್ಚಿತ್ ಮಾಡ್‌ನದ್.
ಅಡಸು (ಕ್ರಿ): press firmly; squeeze: ಜೋರ್ತೆ ಒತ್ತ; ಅವುಂಕ್; ಅಮ್‌ಕ್.	ಅಡಿಗಿ (ನಾ): cooking: ಅಡಿಗಿ; ಅಡ್‌ಗೆ.
ಅಡಾವುಡಿ ¹ (ನಾ): hurry; haste: ಆಶುರ; ಅಲ್‌ಂಬಲ್.	ಅಡಿಗಿಮನೆ (ನಾ): kitchen: ಅಡ್‌ಗೆ ಮನೆ.
ಅಡಾವುಡಿ ² (ನಾ): din; bustle; turmoil: ಅಡಾವುಡಿ; ಗಲಾಟೆ.	ಅಡಿಗಿಮನೆ ಅಟ್ಟಣಿಗೆ (ನಾ): kitchen loft: ಬೆದ್‌ಕೆ; ಬೇಕೆ; ಬೇತೆ.
ಅಡಿ ¹ (ನಾ): foot: ಕಾಲ್‌ರ ಕೇದತ್ ಬಾಗ; ಅಂಗಾಲ್, ಮೇಂಗಾಲ್ ದಂಡೂ ಕೂಡಿತ್.	ಅಡಿಗಿಯವನು (ನಾ): cook: ಅಡ್‌ಗೆಕಾರ್; ಅಡ್‌ಗೆ ಮಾಡ್‌ವಂವ.
	ಅಡಿಗಿ ಸೋಡ (ನಾ): sodium bicarbonate: ಸೋಡ ಪೊಡಿ.
	ಅಡಿಪಾಯ (ನಾ): foundation; base-

Fig 7: Usage of Diacritics. (Kuttappa 2024:18)

ಅರೆ² (ಕ್ರಿ) ಅರ್ತಿತ್ ತುಂಡ್ ಮಾಡ್; ಕತ್ತರಿಚಿಡ್; ನೆರ್ಕ್. sever; cut. ಮುರಿ, ಕುಯ್ಯು.

ಅರೆ³ (ನಾ) ನಡು; ತಡಿರ ನಡುಬರಿ; ಕುಂಡಿರ ಕೂಡಿಕ್ ಕೆಲತಾರ ಕೇದ ಉಳ್ಳ ಬಾಗ. waist. ಸೊಂಟ.

ಅರೆ⁴ (ಕ್ರಿ) ಅರ್ತಾರ್; ತಯಿ; ಕಲ್ಲರ ಕೂಡಿಲ್ ಟ್ಟಂಡ್ ತೇಯಿ. grind. ತೇಯು.

ಅರೆ⁵ (ನಾ) ಕೋಂಬರ ಎಳರೆರ ಒಳ್ಲ್. room. ಕೋಣೆ.

ಅರೆ ಆಲ್ (ನಾ) ಕುಯ್ಯ ಆಲ್ಮರ; ಚಿರಿಯ ಆಲ್ಮರ, ಇಯ್ಯೂ ತೇಂದ ಆಲ್ಮರ ಆಕತದ್. a young *Ficus religiosa* L. tree. ಇನ್ನೂ ಪೂರ್ಣವಾಗಿ ಬೆಳೆಯದ ಅರಳಿಮರ.

ಅರೆ ಚಂಪೋಲೆ (ನಾ) ಮಂಗಲಕಾರ ಚೇಲೆರ ಕೂಡೆ ನಡುಕ್ ಕೆಣಿಪ ಸರ್ಪಣಿ. waist - chain worn by a bridegroom. ಮದುಮಗ ಧರಿಸುವ ಸೊಂಟದ ಸರಪಳಿ.

ಅರೆ ಚಟ್ಟ್ (ನಾ) ಪಾದಿ ಪೊಂಜಿ; ಬೊಳ್ಂದಿತ್ ನಲ್ಲೋಣೆ ತಿರ್ ಕೂಡತ ಕುರು. fully grown but half-developed grain. ಅರ್ಧ ತುಂಬಿದ ಬೆಳೆದ ಕಾಳು.

ಅರೆ ಬಾಚೆಕಾರ (ನಾ) ಅರೆ ಆಯ್ಸತ್ ಅರ ಬಾಚೆಕಾರ - ಪಂಡ್ ಕೊಡಗ್ನ ಆಳಿಯಂಡಿಂಜ ಬಾಚೆಕಾರ ತಂಗಡೊಳ್ಲೇ ಪಡೆ ಪೊತ್ತಂಡ್ ಚನ್ನಂಗ್ ಕಾಲೇ ಬಾಳುವ ಇಂಜಾಂಗ್ ಅಯಿಂಗಳ ಅರೆ ಬಾಚೆ ಕಾರಾಂದ್ ಕಾಕ್ವ ಇಂಜತ್. petty rulers who fought amongst themselves and ruled for a few years only. ಒಬ್ಬರ ಮೇಲೊಬ್ಬರು ಬಿದ್ದು ಕೆಲವೇ ಕಾಲ ಕೊಡಗಿನ ನಾಡುಗಳನ್ನು ಆಳುತ್ತಿದ್ದ ನಾಯಕರು.

ಅರೆಂಜ ಕುತ್ತ (ನಾ) ನಡುರ ಕೇದಕ್; ತಡಿರ ನಡುಬರಿಂಜ ತವ್ವಕ್. below the waist. ಸೊಂಟದ ಕೆಳಗೆ.

ಅರೆಂದಳಿ (ನಾ) ಕೇಂದಳಿ; ಬಾಯಿನ ಮುಚ್ಚಿತ್ ಕಿರಿನ ಮಾತ್ರ ಚನ್ನ ಬೇತಿತ್ ತಳಿಪ; ಚಿರಾಂಬನೆ ಕುಸಿನ ಕಾಟ್ವ. a smile. ಕಿರುನೆಗೆ.

ಅರೆಕಟ್ಟ್ (ನಾ) ಅರ್ಕಟ್ಟ್ - ನೋಟ.

ಅರೆಕಟ್ಟ್ ಚೇಲೆ (ನಾ) ಚೆಂಡೆ ಚೇಲೆ - ನೋಟ.

ಅರೆಪಟ್ಟಿ (ನಾ) ನಡುಪಟ್ಟಿ; ಪೊಣ್ಣಾಲ್ ತಡಿರ ನಡುಕ್ ಕಟ್ಟಿಪ ಪಟ್ಟಿ. waist-band. ಹೆಂಗಸರು ಸೊಂಟದ ಸುತ್ತು ಕಟ್ಟಿಪ ಪಟ್ಟಿ.

ಅರೆಪಡಿಪ್ಕಾರ (ನಾ) ಒರ್ ವಿಸಯತ್ನ ವಿವರೋಡೆ ಅರಿಯತಂಪ; ಚನ್ನಚನ್ನ ಪಡ್ಚಂಪ; ಪಾದಿ ಪಡ್ಚಂಪ. one with half knowledge. ಒಂದು ವಿಚಾರದ ಪೂರ್ಣ ಅರಿವು ಇಲ್ಲದವನು; ಯಾವುದೇ ವಿಚಾರವನ್ನು ಸರಿಯಾಗಿ ಕಲಿತಿಲ್ಲದವನು.

ಅರೆಪಾಟ್ಕಾರ (ನಾ) ಸರಿಯಾಯಿತ್ ಇಡೀ ಪಾಟ್ ಗೊತ್ತಿಲ್ಲತಂಪ; ಅರೆಪಡಿಪ್ಕಾರನ ಎಣ್ಣೆವದ್. one who does not know the complete ballad or song; term used to refer to one with half knowledge. ಪೂರ್ತಿ ಹಾಡು ಗೊತ್ತಿಲ್ಲದ ವನು; ಪೂರ್ಣ ಅರಿವಿಲ್ಲದವನನ್ನು ಹೇಳುವುದು.

ಅರೆಪಿರಾಂದ್ (ನಾ) ಚನ್ನ ಪುಚ್ಚಿಪುಚ್ಚಿ ಕಳಪದ್; ತಕ್ಕಬಾಕ್ ಕಳತಳಿ ಸರಿ ಇಲ್ಲತೆ. ಪೋದ ಇಲ್ಲತವಂಡನ್ನಕೆ ಕಳಪದ್; ಪಿರಾಂದ್ ವೊಳಿ. foolish behaviour. ಮರಳುತನ; ಹುಚ್ಚಾಟಕೆ; ತಿಳಿಗೇಡಿತನ.

ಅರೆಬಟ್ಟಿ (ನಾ) ನಾಲ್ ಪರೆ ನೆಲ್ ಪುಡಿಪನ್ನತ ಪುಂಡ ಇಲ್ಲೆಂಗಿ ವಾಟಿರ ಬಟ್ಟಿ. basket which holds four measures each

Fig 8: Usage of diacritics. Noted: vertical bar is slightly bend. (Uthaiiah & Thangamma 2017:40)

**PROPOSAL SUMMARY FORM TO ACCOMPANY SUBMISSIONS
FOR ADDITIONS TO THE REPERTOIRE OF ISO/IEC 10646¹**

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:	Proposal to Encode the Four Diacritic Marks of Kannada in Unicode	
2. Requester's name:	Biswajit Mandal (biswajitmandal.bm90@gmail.com)	
3. Requester type (Member body/Liaison/Individual contribution):	Individual Member	
4. Submission date:	28/03/2026	
5. Requester's reference (if applicable):		
6. Choose one of the following:		
This is a complete proposal:	Yes	
(or) More information will be provided later:	No	

B. Technical – General

1. Choose one of the following:		
a. This proposal is for a new script (set of characters):	Yes	
Proposed name of script:	Kannada	
b. The proposal is for addition of character(s) to an existing block:		
Name of the existing block:		
2. Number of characters in proposal:	4	
3. Proposed category (select one from below - see section 2.2 of P&P document):		
A-Contemporary <input checked="" type="checkbox"/>	B.1-Specialized (small collection) <input 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?	Yes	
a. If YES, are the names in accordance with the "character naming guidelines" in Annex L of P&P document?	Yes	
b. Are the character shapes attached in a legible form suitable for review?	Yes	
5. Fonts related:		
a. Who will provide the appropriate computerized font to the Project Editor of 10646 for publishing the standard?	Biswajit Mandal (biswajitmandal.bm90@gmail.com)	
b. Identify the party granting a license for use of the font by the editors (include address, e-mail, ftp-site, etc.):	Biswajit Mandal (biswajitmandal.bm90@gmail.com)	
6. References:		
a. Are references (to other character sets, dictionaries, descriptive texts etc.) provided?	Yes	
b. Are published examples of use (such as samples from newspapers, magazines, or other sources) of proposed characters attached?	Yes	
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)?	Yes	
	See proposal for additional details	

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		No
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>Kodava community of India</i> <i>See citations in text proposal</i>	Yes
3. Information on the user community for the proposed characters (for example: size, demographics, information technology use, or publishing use) is included? Reference:	<i>See text of proposal</i>	Yes
4. The context of use for the proposed characters (type of use; common or rare) Reference:	<i>See text of proposal</i>	Common
5. Are the proposed characters in current use by the user community? If YES, where? Reference:	<i>See text of proposal</i>	Yes
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:		No
7. Should the proposed characters be kept together in a contiguous range (rather than being scattered)?		N/A
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:		No
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:		No
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>See text of proposal</i>	Yes Yes
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:		No
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)		No
13. Does the proposal contain any Ideographic compatibility characters? If YES, are the equivalent corresponding unified ideographic characters identified? If YES, reference:		No

UnicodeData.txt

```
1 0CD0 ; KANNADA SIGN OVERLINE;Mn;0;NSM;;;;;N;;;;;
2 0CD1 ; KANNADA SIGN VERTICAL LINE ABOVE;Mn;0;NSM;;;;;N;;;;;
3 0CD2 ; KANNADA SIGN DOT ABOVE;Mn;0;NSM;;;;;N;;;;;
4 0CD3 ; KANNADA SIGN CANDRA ABOVE;Mn;0;NSM;;;;;N;;;;;
```

+4

IndicPositionalCategory.txt

```
0CD0..0CD3; Top # Mn [01] KANNADA SIGN OVERLINE .. KANNADA SIGN CANDRA ABOVE
```

IndicSyllabicCategory.txt

```
0CD0..0CD3; Vowel_Dependent # Mn [04] KANNADA SIGN OVERLINE .. KANNADA SIGN  
CANDRA ABOVE
```