

# Unicode request for compound tone diacritics II L2/23-

Kirk Miller, kirkmiller, gmail.com

2023 September 18

This request is for ligatures, of three tone diacritics, that include the vertical line  $\overset{\cdot}{\text{v}}$  or the macron  $\bar{\text{v}}$ . It follows on L2/23-188, which covered simpler ligatures of the vertical line.

Transcription systems for languages with four distinct pitches commonly use a vertical line for the higher mid pitch than the macron, thus contrasting  $\overset{\cdot}{\text{v}}$  *high*,  $\overset{\cdot}{\text{v}}$  *high-mid*,  $\bar{\text{v}}$  *low-mid*,  $\underset{\cdot}{\text{v}}$  *low*, equivalent to  $\uparrow\uparrow\downarrow$  in Chao tone letters. This occurs for example in descriptions of Kru and Otomanguean languages, and with the values of  $\overset{\cdot}{\text{v}}$  and  $\bar{\text{v}}$  reversed in the General Alphabet of Cameroon Languages. A vertical line may also be used for a single mid tone when the macron is used instead for vowel length. When transcribing contour tones, these diacritics form ligatures analogous to those in the Combining Diacritical Marks Supplement and Combining Diacritical Marks Extended.

A large number of three-pitch contour tone ligatures are theoretically possible. Beside the two already supported by Unicode (U+1DC8  $\overset{\cdot}{\text{v}}$  and U+1DC9  $\underset{\cdot}{\text{v}}$ ), I've been able to attest to three: vertical-acute-grave  $\overset{\cdot}{\text{v}}\grave{\text{v}}$ , vertical-grave-acute  $\grave{\text{v}}\overset{\cdot}{\text{v}}$ , and macron-acute-grave  $\overset{\cdot}{\text{v}}\grave{\text{v}}$ , corresponding to  $\uparrow\downarrow\downarrow$  in Chao tone letters. These are requested in this proposal.

## Characters

- $\overset{\cdot}{\text{v}}\grave{\text{v}}$  1AD6 COMBINING VERTICAL-LINE-ACUTE-GRAVE.
- $\grave{\text{v}}\overset{\cdot}{\text{v}}$  1AD7 COMBINING VERTICAL-LINE-GRAVE-ACUTE.
- $\overset{\cdot}{\text{v}}\grave{\text{v}}$  1AD8 COMBINING MACRON-ACUTE-GRAVE.

## Properties

1AD6;COMBINING VERTICAL-LINE-ACUTE-GRAVE;Mn;230;NSM;;;;;N;;;;;  
 1AD7;COMBINING VERTICAL-LINE-GRAVE-ACUTE;Mn;230;NSM;;;;;N;;;;;  
 1AD8;COMBINING MACRON-ACUTE-GRAVE;Mn;230;NSM;;;;;N;;;;;

# Chart

Greyed-out cells have already been assigned or (light grey) are preliminarily accepted from other proposals.

## Combining Diacritical Marks Extended

1AB0

1ADF

	1AB	1AC	1AD	1AE
0	◌̂	◌̂	◌̂	
1	◌̇	◌̇	◌̇	
2	◌̈	◌̈	◌̈	
3	◌̉	◌̉	◌̉	
4	◌̊	◌̊	◌̊	
5	◌̋	◌̋	◌̋	
6	◌̌	◌̌	◌̌	
7	◌̍	◌̍	◌̍	
8	◌̎	◌̎	◌̎	
9	◌̏	◌̏	◌̏	
A	◌̐	◌̐	◌̐	
B	◌̑	◌̑	◌̑	
C	◌̒	◌̒	◌̒	
D	◌̓	◌̓	◌̓	
E	◌̔	◌̔		
F	◌̕	◌̕		

# References

- Marchese, Lynell (1983) Listes de lexemes Kru. In *Atlas linguistique Kru: essai de typologie*, 323–405. 3rd edition, Abidjan.
- Vogler, Pierre (1988) *Le Parler Vata*, Strasbourg.

# Figures

2.75. Le système tonal se résume comme suit :

	ponctuels		modulés unidirectionnels		modulés bidirectionnels	
registres		descendants	montants	descendants	montants	descendants
haut	/	^		~		
moyen haut		^ ◀	^ ◀	~ ◀	~ ◀	~ ◀
moyen bas	-	^	^			~ ◀
bas	\		v			

Figure 1. Vogler (1988: 52). The complex ligatures marked in purple, <ö, ô, õ>, are requested here. Those marked in blue, <ô, ô>, were requested in part I of this proposal. Unmarked diacritics are already supported by Unicode.

L'ordre de succession des phonèmes et tons est le suivant :

a, ā, b, B, c, d, e, ε, ē, f, g, gb, γ, i, ĩ, t, j, k, kp, l, m, n, ny, η, o, ω, >, ð, p, s, t, u, ū, v, z ; /, |, -, \, ^, v, ^, ^, ^, ^.


Figure 2. Vogler (1988: 421). List of phonemes of Vata. The three-pitch tone <ö> shown here is lexical; the other two in Figure 1 are derivational or intonational.

wobé  
 ï bō à bliā^ ◀ Ne nous frappez pas  
 tu NEG nous frapper

Figure 3. Marchese (1983: 278). Typewriter substitution for the same three-pitch tone diacritic <ö>, from Wobe.

re à celle qui résulte de la concaténation de deux voyelles intonées.

2.68. Une catégorie de modulés bidirectionnels, très rares dans le discours, résulte de la réduction du fonctionnel mlí "dans , à" postposé aux locatifs. Le ton haut provenant de la réduction peut se combiner, sans allongement vocalique perceptible, avec un ton descendant haut ou moyen haut :


  
 n̄́ "à la mère"
   
 Bù "dans le trou"
   
 ébíjā "à Abidjan"

2.69. Les combinaisons ʌ et ʌ ont été observées dans quelques termes non analysables :

k̀̀k̀̀l̀̀k̀̀k̀̀ ʌ "cocorico"
   
 m̄́m̄́ "charognard"
   
 t̄́t̄́ "gouttière"

2.70. L'amalgame de ʌ (<mlí) au ton final bidirectionnel provoque la for-

Figure 4. Vogler (1988: 48). Three-pitch tones <̂, ̃> that result from the conflation of simpler underlying tones in Vata (§2.68), and two more, <̂, ̃>, found in unanalyzable lexical items (§2.69). The diacritic marked in yellow, <̂>, is supported at U+1DC9 COMBINING ACUTE-GRAVE-ACUTE.

k̀̀ j̄́l̄́	: V "se relever"
k̀̀ kl̄́l̄́l̄́l̄́	: V "attraper"
k̀̀kl̄́	: N sg (-kl̄́) "
k̀̀k̀̀k̀̀	: N sg (k̀̀k̀̀k̀̀)
k̀̀k̀̀k̀̀k̀̀ ʌ	: Ph "cocorico"
k̀̀k̀̀k̀̀ pī́	: V "être mal,

Figure 5. Vogler (1988: 460). A three-pitch compound diacritic <̂> containing a vertical line.

p̄́t̄́	: V "aplatir"
p̄́t̄́t̄́	: N sg (-t̄́) "charpente"

Figure 6. Vogler (1988: 486). A three-pitch compound diacritic <̂> containing a macron, contrasting with the vertical line in <̂> in Figure 5.

109 B̄́t̄́ m̄́ é̄́ fl̄́ns̄́ ná kw̄́ B̄́B̄́ w̄́ l̄́ d̄́ḡ́nyw̄́ prends-le en considération.

Figure 7. Vogler (1988: 419). A third three-pitch tone diacritic, <̂>, in running text.

ISO/IEC JTC 1/SC 2/WG 2  
**PROPOSAL SUMMARY FORM TO ACCOMPANY SUBMISSIONS  
 FOR ADDITIONS TO THE REPERTOIRE OF ISO/IEC 10646<sup>1</sup>.**

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

Please read Principles and Procedures Document (P & P) from [std.dkuug.dk/JTC1/SC2/WG2/docs/principles.html](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 [std.dkuug.dk/JTC1/SC2/WG2/docs/summaryform.html](http://std.dkuug.dk/JTC1/SC2/WG2/docs/summaryform.html).  
 See also [std.dkuug.dk/JTC1/SC2/WG2/docs/roadmaps.html](http://std.dkuug.dk/JTC1/SC2/WG2/docs/roadmaps.html) for latest *Roadmaps*.

**A. Administrative**

1. **Title:** Compound tone diacritics II

2. Requester's name: Kirk Miller

3. Requester type (Member body/Liaison/Individual contribution): individual

4. Submission date: 2023 September 18

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:  \_\_\_\_\_

**B. Technical – General**

1. Choose one of the following:

a. This proposal is for a new script (set of characters):  no  
 Proposed name of script: \_\_\_\_\_

b. The proposal is for addition of character(s) to an existing block:  yes  
 Name of the existing block: Combining Diacritical Marks Extended

2. Number of characters in proposal: 3

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? Kirk Miller

b. Identify the party granting a license for use of the font by the editors (include address, e-mail, ftp-site, etc.): SIL (Gentium Release)

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)?  no

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 [www.unicode.org](http://www.unicode.org) for such information on other scripts. Also see Unicode Character Database ([www.unicode.org/reports/tr44/](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.

<sup>1</sup> Form number: N4502-F (Original 1994-10-14; Revised 1995-01, 1995-04, 1996-04, 1996-08, 1999-03, 2001-05, 2001-09, 2003-11, 2005-01, 2005-09, 2005-10, 2007-03, 2008-05, 2009-11, 2011-03, 2012-01)

**C. Technical - Justification**

1. Has this proposal for addition of character(s) been submitted before? If YES explain	<i>no</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>yes</i> <i>Author is a member of the user community.</i>
3. Information on the user community for the proposed characters (for example: size, demographics, information technology use, or publishing use) is included? Reference:	
4. The context of use for the proposed characters (type of use; common or rare) Reference:	<i>transcription</i>
5. Are the proposed characters in current use by the user community? If YES, where? Reference:	<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> <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>