

Unicode request for compound tone diacritics III

Kirk Miller, kirkmiller, gmail.com

2024 October 18

This request is for compounds of the acute and grave accents with a tittle and of caron-acute.

Transcription systems for three distinct pitches commonly use a vertical line for the mid pitch, thus contrasting \acute{o} *high*, $\overset{\cdot}{o}$ *mid*, \grave{o} *low*. Compounds of these diacritics indicate contour tones: rising $\overset{\cdot}{\acute{o}}$ and falling $\acute{o}\grave{o}$. To indicate intermediate pitches, a dot may be added to an acute or grave, for $\overset{\cdot}{\acute{o}}$ *high*, $\overset{\cdot}{\grave{o}}$ *mid-high*, $\overset{\cdot}{\overset{\cdot}{o}}$ *mid*, $\overset{\cdot}{\grave{o}}$ *mid-low*, \grave{o} *low*, equivalent to the five basic Chao tone letters $\langle \uparrow \uparrow \uparrow \downarrow \downarrow \rangle$

The five IPA tone diacritics $\langle \overset{\cdot}{\acute{o}} \overset{\cdot}{\grave{o}} \bar{o} \grave{o} \overset{\cdot}{\grave{o}} \rangle$ are not equivalent to the five basic Chao tone letters. As summarized in the report to the Kiel Convention in 1989, the double diacritics $\langle \overset{\cdot}{\grave{o}} \overset{\cdot}{\grave{o}} \rangle$ are *extra-high* and *extra-low*, whereas the Chao tone letters $\langle \uparrow \downarrow \rangle$ are equivalent to $\langle \overset{\cdot}{\acute{o}} \rangle$ *high* and $\langle \grave{o} \rangle$ *low*, and the Chao tone letters $\langle \uparrow \downarrow \rangle$ intermediate between *high/low* and *mid*. The proposed para-IPA diacritics $\langle \overset{\cdot}{\grave{o}} \overset{\cdot}{\grave{o}} \rangle$ correspond to the intermediate Chao letters $\langle \uparrow \downarrow \rangle$.

The Medieval Unicode Font Initiative (MUFI) supports twenty vowel letters with the dot and acute accent, specifically $\langle \acute{a} \acute{e} \acute{i} \acute{o} \acute{u} \acute{y} \grave{a} \grave{e} \grave{i} \grave{o} \grave{u} \grave{y} \overset{\cdot}{a} \overset{\cdot}{e} \overset{\cdot}{i} \overset{\cdot}{o} \overset{\cdot}{u} \overset{\cdot}{y} \rangle$ and their capitals, in their Private Use Area 32. Both $\overset{\cdot}{o}$ and $\overset{\cdot}{\acute{o}}$ are supported by TIPA.

Descriptions of Khoe languages may distinguish $\overset{\cdot}{\acute{o}}$, $\overset{\cdot}{\grave{o}}$ for high vs high-rising tones, and $\overset{\cdot}{\grave{o}}$, $\overset{\cdot}{\overset{\cdot}{o}}$ for contrasting low-rising tones. The latter consist of low $\langle \grave{o} \rangle$ plus high $\langle \overset{\cdot}{\acute{o}} \rangle$ and low $\langle \grave{o} \rangle$ plus extra-high $\langle \overset{\cdot}{\overset{\cdot}{o}} \rangle$. The distinction is not so much phonetic as a way of marking lexical sets with different tone-sandhi patterns. We also find $\langle \overset{\cdot}{\overset{\cdot}{o}} \rangle$ (mid $\langle \overset{\cdot}{\acute{o}} \rangle$ plus high $\langle \overset{\cdot}{\overset{\cdot}{o}} \rangle$) contrasting with $\langle \overset{\cdot}{\overset{\cdot}{o}} \rangle$ (mid $\langle \overset{\cdot}{\acute{o}} \rangle$ plus extra-high $\langle \overset{\cdot}{\overset{\cdot}{o}} \rangle$). Caron-acute $\langle \overset{\cdot}{\grave{o}} \rangle$ and vertical line-double acute $\langle \overset{\cdot}{\overset{\cdot}{o}} \rangle$ are missing from Unicode and are requested here. Double grave and double acute below, $\langle \grave{\grave{o}} \rangle$ and $\langle \overset{\cdot}{\overset{\cdot}{o}} \rangle$, contrast with double grave and acute accents above the letter.

Characters

- $\overset{\cdot}{\grave{o}}$ 1ADE COMBINING GRAVE-DOT. Figure 1.
- $\overset{\cdot}{\acute{o}}$ 1ADF COMBINING DOT-ACUTE. Figure 2 ff.
- $\overset{\cdot}{\grave{o}}$ 1AEC COMBINING CARON-ACUTE. Figure 4 ff.
- $\overset{\cdot}{\overset{\cdot}{o}}$ 1AED COMBINING VERTICAL-LINE-DOUBLE-ACUTE. Figure 6 ff.
- $\grave{\grave{o}}$ 1AEE COMBINING DOUBLE GRAVE ACCENT BELOW. Figure 10 ff.
- $\overset{\cdot}{\overset{\cdot}{o}}$ 1AEF COMBINING DOUBLE ACUTE ACCENT BELOW. Figure 11.
- $\overset{\cdot}{\overset{\cdot}{o}}$ 1AF0 COMBINING DOUBLE COMMA ABOVE. Figure 12 ff.

Properties

1ADE;COMBINING GRAVE-DOT;Mn;230;NSM;;;;;N;;;;;
1ADF;COMBINING DOT-ACUTE;Mn;230;NSM;;;;;N;;;;;
1AEC;COMBINING CARON-ACUTE;Mn;230;NSM;;;;;N;;;;;
1AED;COMBINING VERTICAL-LINE-DOUBLE-ACUTE;Mn;230;NSM;;;;;N;;;;;
1AEE;COMBINING DOUBLE GRAVE ACCENT BELOW;Mn;220;NSM;;;;;N;;;;;
1AEF;COMBINING DOUBLE ACUTE ACCENT BELOW;Mn;220;NSM;;;;;N;;;;;
1AF0;COMBINING DOUBLE COMMA ABOVE;Mn;230;NSM;;;;;N;;;;;

Annotations

1AF0 COMBINING DOUBLE COMMA ABOVE
→ 0313 COMBINING COMMA ABOVE
→ 02EE MODIFIER LETTER DOUBLE APOSTROPHE

References

- Crazzolara, Pasquale J. (1960) *A study of the Logbara (Ma'di) language: grammar and vocabulary*. London & New York: Oxford Univ. Press.
- Kilian-Hatz, Christa (2008) *A Grammar of Modern Khwe*. Rüdiger Köppe Verlag, Köln.
- Köhler, Oswin (1981) La langue Kxoe. In J. Perrot (ed.) *Les Langues dans le Monde Ancien et Moderne*. Centre National de la Recherche Scientifique, Paris, pp. 485–555.
- (1989) *Die Welt der Khoé-Buschleute im südlichen Afrika*. Vol I: *Die Khoé-Buschleute und ihre ethnische Umbeugung*. Dietrich Reimer.
- LAMSAS archives, Linguistic Atlas Project, University of Georgia.
[www.lap.uga.edu/Projects/LAMSAS/Pages/LAMSAS Page 1/](http://www.lap.uga.edu/Projects/LAMSAS/Pages/LAMSAS%20Page%201/)
- Navarro Tomás, Tomás. 1962. *Atlas lingüístico de la Península Ibérica (ALPI)*, tomo 1 “Fonética”. Madrid: Consejo superior de investigaciones científicas. <http://alpi.csic.es/en/>.

Chart

Greyed-out cells have already been assigned or (light grey) are requested in other proposals.

Combining Diacritical Marks Extended

1AB0

1ADF

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

Figures

The *low intermediate* (2/5) tone is employed in the following cases :

(a) All dissyllabic personal pronouns have a low-tone first syllable and a slightly higher (2/5) second one, ²
èrì *he . . .* **àmà** *we . . .* **èmì** *you . . .* **èì** *they, them*

(b) On the relative element **-rì**, when suffixed to low-tone syllables, as in the case of the demonstrative pronouns :
'dìrì / **'dìì** *this* **'dàrì** / **'dàì** *that*

(c) On the second syllable of **àlò** *one*

Figure 1. Crazzolara (1960: 11). Grave-dot, <¨>, in Logbara. The numbering convention “2/5” is equivalent to the Chao tone letter <˩>. <ì> is IPA [i].

57. The *high-intermediate* or *raised mid* tone ⁴(4/5) is found mainly in the following cases :

(a) In the full-form personal pronouns (1st and 2nd pers. sg.), i.e. the monosyllables **mǎ** *I*, **mí** *thou*, **'í** *he, she, him- or her-self*.

(b) A good proportion of nouns of kinship (and a few others) end in a high plus an additional raised mid- (4/5) toned syllable. The latter is apparently the result of a shortening of an end syllable **-pǐ** to **-í**, or **-ípǐ** to **-íí**. Examples :
ándrǐí *mother* **átǐí** *father* **ádrǐí** *brother*
ámvíí *sister* **ágúí** or **ágíí** *friend* **'ípǐ** or **'íí** *owner*

(c) There are a number of dissyllabic nouns which in some parts of the country are pronounced with level high tone, while in other parts their first syllable has a high intermediate (4/5) tone. This latter feature is, however, apparently on the point of giving way to the former or high-level tone pronunciation. Examples :
pǎtí *tree* **átǐí** *father* **émvó** *pot*
cf. pǎtí mà pátí *stem of a tree*

Figure 2. Crazzolara (1960: 11–12). Dot-acute, <˙>, in Logbara. The numbering convention “4/5” is equivalent to the Chao tone letter <˨˥>.

Ä	00EBF4	Ä	LATIN CAPITAL LETTER A WITH DOT ABOVE AND ACUTE &Adotacute; (MUFI)
ä	00EBF5	ä	LATIN SMALL LETTER A WITH DOT ABOVE AND ACUTE &adotacute; (MUFI)
Ï	00EBF6	Ï	LATIN CAPITAL LETTER I WITH DOT ABOVE AND ACUTE &ldotacute; (MUFI)
ï	00EBF7	ï	LATIN SMALL LETTER I WITH DOT ABOVE AND ACUTE &idotacute; (MUFI)
Ö	00EBF8	Ö	LATIN CAPITAL LETTER O WITH DOT ABOVE AND ACUTE &Odotacute; (MUFI)
ö	00EBF9	ö	LATIN SMALL LETTER O WITH DOT ABOVE AND ACUTE &odotacute; (MUFI)
Ï	00EBFC	Ï	LATIN CAPITAL LETTER O WITH STROKE AND DOT ABOVE AND ACUTE Ødotacute; (MUFI)
ö	00EBFD	ö	LATIN SMALL LETTER O WITH STROKE AND DOT ABOVE AND ACUTE ødotacute; (MUFI)
Û	00EBFE	Û	LATIN CAPITAL LETTER U WITH DOT ABOVE AND ACUTE &Udotacute; (MUFI)
ú	00EBFF	ú	LATIN SMALL LETTER U WITH DOT ABOVE AND ACUTE &udotacute; (MUFI)

Figure 3. Some of the MUFI vowel letters with dot and acute accent. From the mufi.info website.

Similarly, he distinguishes two Low-High rising tones by using different tone symbols: **ǎ** vs. **ạ̌**.

Figure 4. Kilian-Hatz (2008: 25).

18.	lOéi-goè	Informant	Dikúndù 1979
19.	Bạ̌	Informant	Dikúndù 1980
20.	lÛ	Informant	Dikúndù 1980
21.	lÛi-goè	Informantin	Dikúndù 1981

Figure 5. Köhler, Oswin (1989: 23). Caron-acute <ạ̌> in a list of personal names. <ạ̌> in line 18 was accepted with a recent proposal.

7 **čěú** hè'è ≠ **ạ̌**. ná = xáo.kò **čạ̌** /ám littéralement : « ayant caché j'arrive à deux ».

Figure 6. Köhler (1981: 545). <ạ̌> and <ạ̣̌> (also <ạ̣̣̌> and <ạ̣̣̣̌>). The letters C c with a left tie, used for pre-palatal affricates, can perhaps be handled with U+1AB9 as <č̣̣>, thus <č̣̣ěú>.

12	.à joncture I	. òè suffixe de présent = est couché, a coutume de	qàráti arbre sp.	.m genre masc. sg.	ǎ feuilles	ki dans, sur (postposition)	
18	ǎ griller	.wà joncture I	. òè suffixe de présent = a coutume	tà être ainsi	=xú suffixe de terminatif (indéfini)	=nò quand	xéri broyer

Figure 7. Köhler (1981: 554). <ǎ> and <ạ̌> in running text, along with <ạ̣̌>. The diacritic under the click letters is a narrow allograph of U+1AB6 <ạ̣̌>; in the explanatory material it appears with the normal three humps.

ligne 1	ạ̣̌ « prendre », dans certain marier », etc. Ici : « se procurer (sort).
1-2	hạ̣̌ « faire » ; à l'intention de la chose ».
3	ōáná « porter » ; très exact parlant d'une seule personne
5	ạ̌ná allomorphe de ạ̌nú
9	qàráti sorte d'arbre ; cf.
14	ạ̌xạ̌ « étendre » (une peau)
18	ạ̌ « griller » « calciner » (un objet qui doit être calciné est p...)
25	ạ̌á ici : « faire pénétrer

Figure 8. Köhler (1981: 555). <ạ̌> and <ạ̣̌> in lexical entries.

<i>hámbé</i> s'emploie dans divers contextes :	
1° ne pas être présent, faire défaut.	
ạ̌'á hámbé	« il n'y a pas d'eau là »
Le verbe <i>ạ̌</i> 'ne pas posséder, ne pas avoir'.	
présent : <i>ti ạ̌ wà.tè</i>	préterit : <i>ti djù.wá.hạ̌</i>
<i>ti djù.wá.hạ̌</i> /'áo 'à	« je n'ai pas » (« ne possède pas ») « d'argent »

Figure 9. Köhler (1981: 537).

Ndó †'oá-kxòe.mà
 †Am.kú kx'úi-tcáo-kxòe.mà
 Ndó: |Ándá.m ɔ́ú kí tcá tǐ yá Djwáeni.m 'à kũ.t.à.
 hí rè kàré?
 †Am: tí kũ.t.à.hí
 Ndó: ||ná-tá tcá kũ=ù.wá.kò yà=xú nò áβùrù.m 'à tcá
 ná-tá hí.t.à.hí?

Figure 10. Köhler (1989: 570). Double grave below <◌◌> in running text, including the name of one of the participants of the conversation.

cìráhà támá.xà Qúwá-||háó támá.xà |x'ân |x'oé.kò
 tɛ́.hí; ||ná-tá áβùrù.m 'à †ù.wá=xú nò kúrí.n 'n ||'áe.
 m kí ɔ́gwá=xú nò |ú.m kà tɛ́ kyáo.xá.kũ.t.à.hí; tà=xú=
 nò kyàré.m kà B ú m á. m ó=kà kũ.wá ||oé-tán.à.kò
 ||'áe.m ó=kà cí.xá.kũ.t.à.hí; tǎ.kò cí cǎ.ná=xú nò
 xǎ.||ú 'm áβùrù.m 'à yá cìréyì.m kà oáná xǎ.kũ.t.à.hí;

Figure 11. Köhler (1989: 551). Double grave below <◌◌> and double acute below <◌◌>. Single grave below also occurs.

f00
 †'a = B3
 See introd. note
 glottal low pitch
 trill

Figure 12. LAMSAS archives (LAMSASDC001XI0575.jpg). Double comma above <◌◌>.

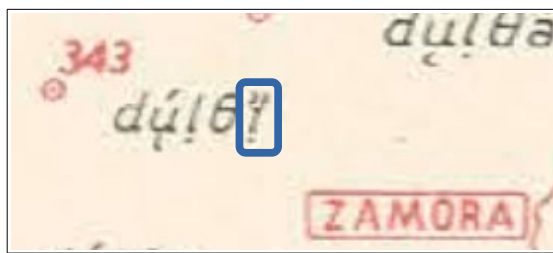


Figure 13. Navarro Tomás (1962: map 74). The double open symbol <̣> is equivalent to a double comma above when the vowel letter is rotated.

ISO/IEC JTC 1/SC 2/WG 2
**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 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.
 See also 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: 2024 October 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: 7

3. Proposed category (select one from below - see section 2.2 of P&P document):

A-Contemporary B.1-Specialized (small collection) B.2-Specialized (large collection)
 C-Major extinct D-Attested extinct E-Minor extinct
 F-Archaic Hieroglyphic or Ideographic G-Obscure or questionable usage symbols

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 for such information on other scripts. Also see Unicode Character Database (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.

¹ 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	<input type="text"/>	<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:	<input type="text"/> <i>Author is a member of the user community.</i>	<i>yes</i>
3. Information on the user community for the proposed characters (for example: size, demographics, information technology use, or publishing use) is included? Reference:	<input type="text"/>	
4. The context of use for the proposed characters (type of use; common or rare) Reference:	<input type="text"/> <i>transcription</i>	
5. Are the proposed characters in current use by the user community? If YES, where? Reference:	<input type="text"/>	<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:	<input type="text"/>	<i>no</i>
7. Should the proposed characters be kept together in a contiguous range (rather than being scattered)?	<input type="text"/>	<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:	<input type="text"/>	<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:	<input type="text"/>	<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:	<input type="text"/>	<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:	<input type="text"/>	<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)	<input type="text"/>	<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:	<input type="text"/>	<i>no</i>