

# Unicode request for stacked arrowheads

Kirk Miller, kirkmiller, gmail.com

2025 June 06

Spacing arrowheads for phonetic notation are found at U+02C2..02C5 and U+02EF..02F2 in the Spacing Modifier Letter block. These arrowheads may be stacked vertically, and the resulting combinations require distinct Unicode characters. Four combinations are used in LAMSAS and are assigned four computer-input macros (Kretzschmar 1993: 143). In publication, the LAMSAS stacked arrowheads consist of a left or right (advanced or retracted) arrowhead over an up or down (raised or lowered) arrowhead, thus: <̛̘̙̚>. Four analogous combinations are used in other Americanist traditions (Figure 4), where the raised/lowered symbol is placed above the advanced/retracted symbol: <̛̘̙̚>. We consider these two arrangements to be allographic, partly because we lack any evidence that they are used contrastively, but also because in the LAMSAS manuscript archive the arrows often appear in the latter arrangement (Figure 5), and were apparently normalized to the former arrangement in publication. A typographer might handle the difference with style variants in the font, if graphic fidelity is determined to be of any importance. It seems unlikely that there could be any semantic distinction between, for example, a *raised retracted* vowel [e̘̙] and a *retracted raised* vowel [e̙̘].

These arrowheads become superscript when they modify superscript letters (Figure 6). It is not clear however that a distinction needs to be made in the encoding, and this should perhaps be left to the font. In better fonts the arrowheads display well in such positions without special formatting; compare Gentium and Noto baseline <ə̘̙ ə̛̚ ə̛̘̙̚>, <ə̛̘̙̚ ə̛̘̙̚> with baseline arrowheads following superscript letters in <ə̛̘̙̚ ə̛̘̙̚>, <ə̛̘̙̚ ə̛̘̙̚>. Even when there is a contrast across a single vowel, as in Figure 7, sequencing may be enough to preserve the underlying text.

Since the characters U+02C2..02C5 are defined as ‘symbol, modifier,’ a possible approach short of separate encoding would be to treat them like U+02DE MODIFIER LETTER RHOTIC HOOK. When the IPA evaluated the draft to L2/20-266 that proposed a complete set of modifier IPA letters, they declined to request modifier variants of U+025A ə̘̙ and U+025D ɜ̘̙, reasoning that the modifier hook should be adequate for any vowel letter, baseline or superscript. At the time, fonts such as Gentium made no special accommodation, yet the resulting sequence <̘̙> was perfectly legible. Since then Gentium has accommodated the relative height of modifier letters so that <̘̙> now displays as <̘̙̘̙>. Such accommodation might be made for the arrowheads as well; if ambiguity arose from simple strings of superscript letters and baseline arrowheads, perhaps a zero-width (non)joiner might be utilized either to force superscripting of the arrowheads or to prevent superscripting. A note could be added to the CoreSpec that this is something typographers might consider. However, a potential complication is that U+02C2..02C5 are not defined as diacritics. Regardless, it would be premature to request distinct characters at this point.

# Characters

## Compound arrowheads

- ⤵ 1DFC9 MODIFIER LETTER LEFT ARROWHEAD OVER UP ARROWHEAD.
- ⤴ 1DFCA MODIFIER LETTER RIGHT ARROWHEAD OVER UP ARROWHEAD.
- ⤶ 1DFCB MODIFIER LETTER LEFT ARROWHEAD OVER DOWN ARROWHEAD.
- ⤷ 1DFCC MODIFIER LETTER RIGHT ARROWHEAD OVER DOWN ARROWHEAD.

## Properties

1DFC9;MODIFIER LETTER LEFT ARROWHEAD OVER UP ARROWHEAD;Sk;0;ON;;;;;N;;;;;  
1DFCA;MODIFIER LETTER RIGHT ARROWHEAD OVER UP ARROWHEAD;Sk;0;ON;;;;;N;;;;;  
1DFCB;MODIFIER LETTER LEFT ARROWHEAD OVER DOWN ARROWHEAD;Sk;0;ON;;;;;N;;;;;  
1DFCC;MODIFIER LETTER RIGHT ARROWHEAD OVER DOWN ARROWHEAD;Sk;0;ON;;;;;N;;;;;

## Annotations

Under the character subheading, a common annotation could be added:

The following characters may have the inverse stacking order of up/down arrowhead over left/right.  
They are baseline characters.

## References

- Carlson (1972) *A Grammar of Spokane*. University of Hawaii, Honolulu.
- Kretzschmar (1993) *Handbook of the Linguistic Atlas of the Middle and South Atlantic States*. University of Chicago Press.
- Larry Mattes & Donald Omark (1991) *Speech and language assessment for the bilingual handicapped*. 2nd edition, College-Hill Press.
- Raven McDavid & Raymond O’Cain (1980) *Linguistic Atlas of the Middle and South Atlantic States*. University of Chicago Press, fasc. 2.

# Charts

## Latin Extended-G

1DF00

1DFFF

	1DF0	1DF1	1DF2	1DF3	1DF4	1DF5	1DF6	1DF7	1DF8	1DF9	1DFA	1DFB	1DFC	1DFD	1DFE	1DFF
0	ṡ	ṡ	ḡ	ḡ	Ḃ	Ḃ	ḡ	ḡ	Ḃ					ḡ	ḡ	Ḃ
1	ṡ	ḡ	ḡ	ḡ	Ḃ	Ḃ	ḡ	ḡ	Ḃ					ḡ	ḡ	Ḃ
2	ḡ	ḡ	ḡ	ḡ	Ḃ	Ḃ	ḡ	ḡ	Ḃ					ḡ	ḡ	Ḃ
3	ḡ	ḡ	ḡ	ḡ	Ḃ	Ḃ	ḡ	ḡ	Ḃ					ḡ	ḡ	Ḃ
4	ḡ	ḡ	ḡ	ḡ	Ḃ	Ḃ	ḡ	ḡ	Ḃ					ḡ	ḡ	Ḃ
5	ḡ	ḡ	ḡ	ḡ	Ḃ	Ḃ	ḡ	ḡ	Ḃ					ḡ	ḡ	Ḃ
6	ḡ	ḡ	ḡ	ḡ	Ḃ	Ḃ	ḡ	ḡ	Ḃ					ḡ	ḡ	Ḃ
7	ḡ	ḡ	ḡ	ḡ	Ḃ	Ḃ	ḡ	ḡ	Ḃ					ḡ	ḡ	Ḃ
8	ḡ	ḡ	ḡ	ḡ	Ḃ	Ḃ	ḡ	ḡ	Ḃ					ḡ	ḡ	Ḃ
9	ḡ	ḡ	ḡ	ḡ	Ḃ	Ḃ	ḡ	ḡ	Ḃ				ḡ	ḡ	Ḃ	Ḃ
A	ḡ	ḡ	ḡ	ḡ	Ḃ	Ḃ	ḡ	ḡ	Ḃ				ḡ	ḡ	Ḃ	Ḃ
B	ḡ	ḡ	ḡ	ḡ	Ḃ	Ḃ	ḡ	ḡ	Ḃ				ḡ	ḡ	Ḃ	Ḃ
C	ḡ	ḡ	ḡ	ḡ	Ḃ	Ḃ	ḡ	ḡ	Ḃ				ḡ	ḡ	Ḃ	Ḃ
D	ḡ	ḡ	ḡ	ḡ	Ḃ	Ḃ	ḡ	ḡ	Ḃ				ḡ	ḡ	Ḃ	Ḃ
E	ḡ	ḡ	ḡ	ḡ	Ḃ	Ḃ	ḡ	ḡ	Ḃ				ḡ	ḡ	Ḃ	Ḃ
F	ḡ	ḡ	ḡ	ḡ	Ḃ	Ḃ	ḡ	ḡ	Ḃ				ḡ	ḡ	Ḃ	Ḃ

## Figures

Double shift signs (one pointing to the left or right, the other up or down) are used to indicate shades of sound for which single shift signs are inadequate. These combinations must be interpreted according to the statements above. Thus [i̯] is a vowel between [i] and [ɨ] and [ɪ] but nearest to [i]; [o̞] is a vowel between [o] and [ə] and [ɔ] but nearest to [o]; [ə̘] is a vowel between [ə] and [ɛ] (or [e]) and [ɨ] but nearest to [ə]; [a̠] is a vowel between [a] and [ɐ] but nearer to [a].

Figure 1. Kretzschmar (1993: 118). The four combinations of left-right arrowheads over up-down arrowheads,  $\langle \nearrow \searrow \swarrow \nwarrow \rangle$ , placed after the vowel.

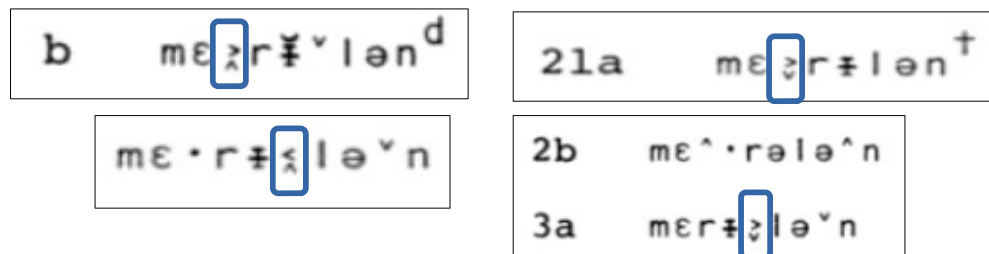


Figure 2. McDavid & O’Cain (1980: 119–124). Textual examples of  $\langle \lambda \ \acute{\lambda} \ \rangle$ .

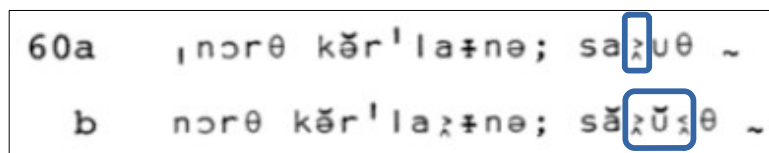


Figure 3. McDavid & O’Cain (1980: 138).  $\langle \dot{\lambda} \rangle$  and  $\langle \dot{\lambda} \rangle$ .

cəʔúk<sup>w</sup>əlt. 'It is brought.'

[ʔ>], central low rounded, in the neighborhood of rounded postvelars.

q<sup>w</sup>ələmín. '(They're) ashes.'

Figure 4. Carlson (1972: 14). An example of <ʔ> preceding a vowel. It has the inverse stacking order of up/down over left/right.

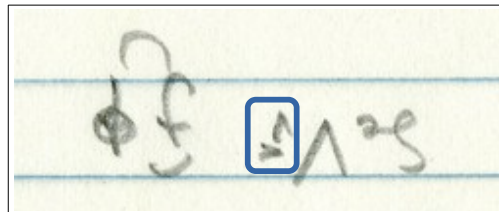


Figure 5. LAMSAS archive, [page 1A VA4N](#). <ʔ> placed to the left of the vowel letter, with the inverse stacking convention of up/down over left/right. The stacking order would be normalized in publication.

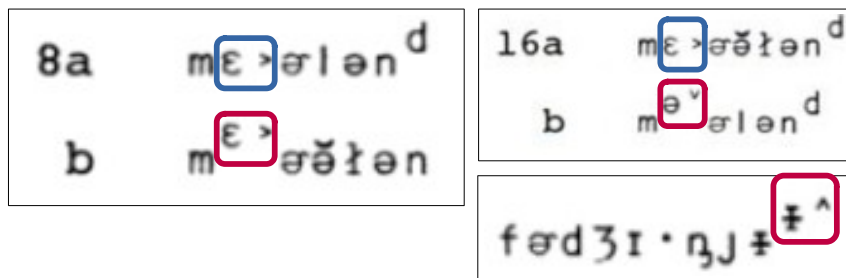


Figure 6. McDavid & O'Cain (1980: 120, 119, 128). When a vowel is made a modifier superscript, the arrow heads are made superscript as well (red).

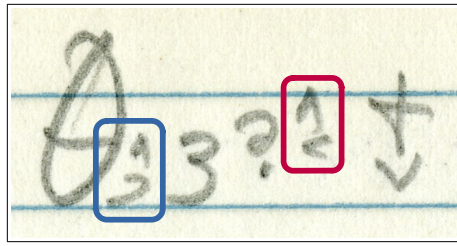


Figure 7. LAMSAS archive, [page 1A VA4N](#). A contrast between baseline and superscript stacked arrows, with <̣> placed to the left of the baseline vowel letter, as in Figure 5 also from this page, and superscript <̣> placed to the right of the superscript schwa <̣>. The sequencing is enough to preserve the distinction, so encoding a superscript <̣> does not seem necessary.

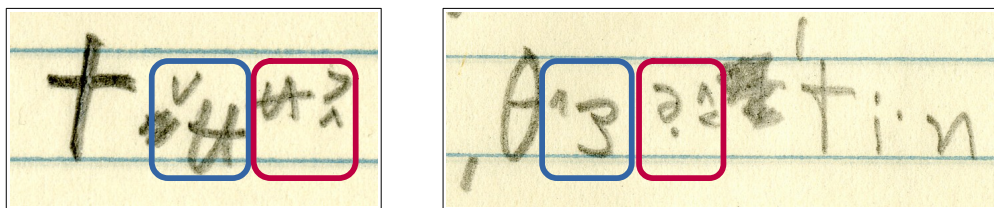


Figure 8. LAMSAS archive, [page 1 VAN4](#). In this manuscript, the arrowheads aren't placed any higher for superscript letters than otherwise. Compare the distinction in print and in the previous figure.

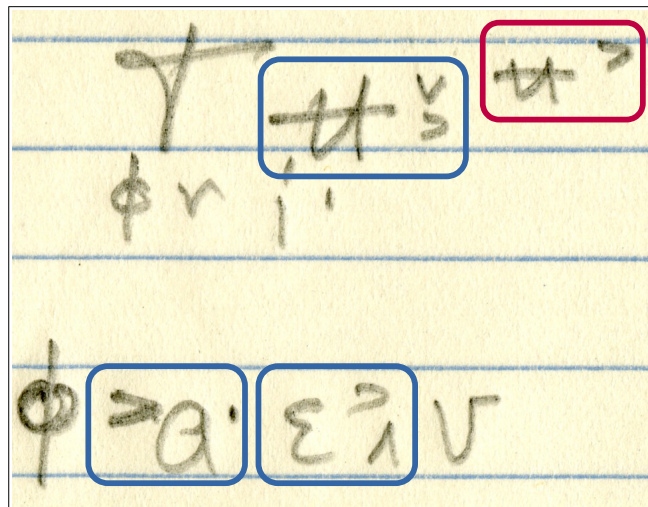


Figure 9. LAMSAS archive, [page 1 VA12](#). At top the arrowheads which modify the baseline letter separate it from three superscript, which has its own arrowhead; at bottom, arrowheads are placed to the left and right of a pair of baseline letters. On page 4 VA43 we find a similar <'ṭaẹṃ> with single arrowheads; similarly elsewhere.

Affricates	p <sup>f</sup>	t <sup>θ</sup>	t <sup>s</sup>	t <sup>s</sup>	t <sup>l</sup>	ç	č	c <sup>f</sup>	k <sup>x</sup>
		d <sup>θ</sup>	d <sup>z</sup>	d <sup>z</sup>	d <sup>l</sup>	ɟ	ɟ	j <sup>j^</sup>	g <sup>l</sup>
Fricatives	ɸ	f	θ	ɬ	s	ʃ	ʃ	ç	x
	β	v	ð	ɮ	z	ʒ	ʒ	j <sup>j^</sup>	ɣ
Approximants	ɹ		ɻ	ɻ			j		

Figure 10. Mattes & Omark (1991: 184). <j^> (raised j) is used for IPA <j> in this Americanist notation. Affricates are indicated by superscripting the fricative, thus <j<sup>i^</sup>> with a superscript arrowhead is needed for IPA <ɟ>. Once again, typefaces such as Gentium and Noto may be adequate for typesetting this notation without dedicated Unicode superscript arrowheads, but it raises the possibility that such characters may be needed in the future.

**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 <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**

<b>1. Title:</b>	<i>Stacked arrowheads</i>
<b>2. Requester's name:</b>	<i>Kirk Miller</i>
<b>3. Requester type (Member body/Liaison/Individual contribution):</b>	<i>individual</i>
<b>4. Submission date:</b>	<i>2025 June 06</i>
<b>5. Requester's reference (if applicable):</b>	
<b>6. Choose one of the following:</b>	
This is a complete proposal:	<i>yes</i>
(or) More information will be provided later:	

**B. Technical – General**

<b>1. Choose one of the following:</b>	
a. This proposal is for a new script (set of characters):	<i>no</i>
Proposed name of script:	
b. The proposal is for addition of character(s) to an existing block:	<i>yes</i>
Name of the existing block: <i>Latin Extended-G</i>	
<b>2. Number of characters in proposal:</b>	<i>4</i>
<b>3. Proposed category (select one from below - see section 2.2 of P&amp;P document):</b>	
A-Contemporary <input type="checkbox"/>	B.1-Specialized (small collection) <input checked="" type="checkbox"/>
C-Major extinct <input type="checkbox"/>	B.2-Specialized (large collection) <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"/>
<b>4. Is a repertoire including character names provided?</b>	<i>yes</i>
a. If YES, are the names in accordance with the “character naming guidelines” in Annex L of P&P document?	
b. Are the character shapes attached in a legible form suitable for review?	
<b>5. Fonts related:</b>	
a. Who will provide the appropriate computerized font to the Project Editor of 10646 for publishing the standard?	
b. Identify the party granting a license for use of the font by the editors (include address, e-mail, ftp-site, etc.):	
<i>SIL (Gentium Release)</i>	
<b>6. References:</b>	
a. Are references (to other character sets, dictionaries, descriptive texts etc.) provided?	<i>yes</i>
b. Are published examples of use (such as samples from newspapers, magazines, or other sources) of proposed characters attached?	<i>yes</i>
<b>7. Special encoding issues:</b>	
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)?	
<i>no</i>	

**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.

<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?	no
If YES explain	
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.)?	yes
If YES, with whom?	
Rees is involved in a digitization project.	
If YES, available relevant documents:	
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)	republication
Reference:	
5. Are the proposed characters in current use by the user community?	no
If YES, where? Reference:	
6. After giving due considerations to the principles in the P&P document must the proposed characters be entirely in the BMP?	no
If YES, is a rationale provided?	
If YES, reference:	
7. Should the proposed characters be kept together in a contiguous range (rather than being scattered)?	no
8. Can any of the proposed characters be considered a presentation form of an existing character or character sequence?	no
If YES, is a rationale for its inclusion provided?	
If YES, reference:	
9. Can any of the proposed characters be encoded using a composed character sequence of either existing characters or other proposed characters?	no
If YES, is a rationale for its inclusion provided?	
If YES, reference:	
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?	no
If YES, is a rationale for its inclusion provided?	
If YES, reference:	
11. Does the proposal include use of combining characters and/or use of composite sequences?	no
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?	no
If YES, reference:	
12. Does the proposal contain characters with any special properties such as control function or similar semantics?	no
If YES, describe in detail (include attachment if necessary)	
13. Does the proposal contain any Ideographic compatibility characters?	no
If YES, are the equivalent corresponding unified ideographic characters identified?	
If YES, reference:	