L2/24-052

Unicode request for modifier pre-Kiel click letters

Kirk Miller, kirkmiller, gmail.com 2024 February 21

This proposal, officially supported by the International Phonetic Association after evaluation by the IPA Alphbets, Charts and Fonts Committee (Nicolaides 2024), follows on L2/20-252 and L2/20-253, which requested modifier (superscript) variants of nearly all letters of the modern IPA alphabet that were as yet unsupported by Unicode. Those two proposals included a few retired IPA letters, but consideration of historical IPA modifiers was largely deferred until the IPA could more fully discuss the desired scope of such encoding through the newly established Alphabet, Charts and Fonts Committee. This proposal is one result of that consultation.

The Kiel Convention of 1989 resulted in an overhaul of the IPA alphabet, and the click letters were mostly replaced. However, the three pre-Kiel click letters ⟨ᶁ ᶝ ᶇ⟩ remain popular in some circles. The associated letter for palatal clicks, ⟨ᶂ⟩, was never officially part of the IPA, but is used by those linguists who use the pre-Kiel IPA letters. The original IPA letter for palatal clicks, ⟨ᶄ⟩ (see *Écriture phonétique internationale*, 1921: 8, where the four letters ⟨ᶁ ᶇ ᶄ ᶝ⟩ are defined as the four click articulations of Khoekhoe), was defined as ‘velar’ in the 1949 *Principles*, following Jones’ analysis, then retired some years later because velar clicks were judged to be impossible. ⟨ᶄ⟩ has however seen a revival as a *rear-released* velar click, which is used para-lexically across West Africa. In addition, Doke, one of the few early phoneticians to document true retroflex clicks, used the Greek letter psi ⟨ᵠ⟩ for that articulation. The pre-Kiel letter for bilabial clicks, ⟨ᶗ⟩, remains in the IPA, and the corresponding modifier letter ⟨ᵒ⟩ is encoded at U+107B5. We request modifier support for the remaining letters.

Some members of the Alphabet, Charts and Fonts Committee of the IPA expressed reservations about the forms of two of the letters in this document: the stretched *c* and the double-barred esh. In IPA material (Figure 5), the stretched *c* often has a straight left side, ⟨ᶝ⟩, rather than the round form ⟨ᶜ⟩ of the Gentium font used to typeset this document. It is also compressed laterally compared to the letter *c*. However, the original design from 1921 (Figure 3) is very close to the Gentium glyph, and the TIPA typeface preferred by the IPA acknowledges both round and straight allographs of ⟨ᶝ⟩ (Figure 7), though the straight TIPA glyph has an ascender that is not attested from pre-Kiel IPA usage. Similarly, the esh in IPA material has a straight vertical stretch, ⟨ᶋ⟩, like a mathematical integral sign, not ‘wavy’ like the Gentium ⟨ᶊ⟩, and it is also compressed laterally. However, Beach, who invented the symbol ⟨ᶂ⟩, used an uncompressed variant quite a bit ‘wavier’ than even the Gentium form, and a stretched *c* very much like the Gentium form (Figure 4). Late use in publication has a straight ⟨ᶝ⟩ but a wavy ⟨ᶂ⟩, neither laterally compressed (Figure 6). Whatever the designs, the modifier glyphs should be consistent with the base letters.

Characters

ᵻ 107BB MODIFIER LETTER SMALL TURNED T (dental, equivalent to post-Kiel ⟨ǀ⟩)

ᶜ 107BC MODIFIER LETTER STRETCHED C (alveolar, equivalent to post-Kiel ⟨ǃ⟩)

ᶚ 107BD MODIFIER LETTER INVERTED GLOTTAL STOP (lateral, equivalent to post-Kiel ⟨ǁ⟩)

ᶿ 107BE MODIFIER LETTER SMALL ESH WITH DOUBLE BAR (palatal, equivalent to post-Kiel ⟨ǂ⟩)

ᶮ 107BF MODIFIER LETTER SMALL TURNED K (velar, no Kiel equivalent)

Deferred

The modifier psi is deferred pending a decision on whether it should be encoded as Latin or Greek, though there is currently no Latin psi in Unicode.

ᵠ MODIFIER LETTER (GREEK) SMALL PSI (retroflex, equivalent to post-Kiel ⟨ᶑ⟩)

Properties

107BB;MODIFIER LETTER SMALL TURNED T;Lm;0;L;<super> 0287;;;;N;;;;;

107BC;MODIFIER LETTER INVERTED GLOTTAL STOP;Lm;0;L;<super> 0296;;;;N;;;;;

107BD;MODIFIER LETTER SMALL ESH WITH DOUBLE BAR;Lm;0;L;<super> 1DF0B;;;;N;;;;;

107BE;MODIFIER LETTER STRETCHED C;Lm;0;L;<super> 0297;;;;N;;;;;

107BF;MODIFIER LETTER SMALL TURNED K;Lm;0;L;<super> 029E;;;;N;;;;;

Annotations

Annotations identifying the proposed letters with their Lepsius (pipe-letter) equivalents would be useful.

107BB MODIFIER LETTER SMALL TURNED T

→ 107B6 MODIFIER LETTER DENTAL CLICK

107BC MODIFIER LETTER INVERTED GLOTTAL STOP

→ 107B7 MODIFIER LETTER LATERAL CLICK

107BD MODIFIER LETTER SMALL ESH WITH DOUBLE BAR

→ 107B8 MODIFIER LETTER ALVEOLAR CLICK

107BE MODIFIER LETTER STRETCHED C

→ A71D MODIFIER LETTER RAISED EXCLAMATION MARK

References

D.M. Beach (1938) *The Phonetics of the Hottentot Language.* W. Heffer & Sons, Cambridge.

Derek Elderkin (1989) *The Significance and Origin of the Use of Pitch in Sandawe*. University of York.

Rei Fukui (n.d.) *TIPA Typefaces*.

International Phonetic Association (1949) *Principles of the International Phonetic Association*. University College, London.

John Kelly & John Local (1989) *Doing Phonology: Observing, Recording, Interpreting.* Manchester University Press.

Katerina Nicolaides (2024) 'Unicode support for historical and para-IPA letters.' Letter submitted to the Unicode Technical Committee, 01 January 2024. L2/24-049.

Paul Passy & Daniel Jones (Association phonétique internationale, 1921, 2nd ed.) *L'Écriture phonétique internationale: exposé populaire avec application au français et à plusieurs autres langues*.

Chart

This proposal will complete the Latin Extended-F block, apart from two reserved code points.

Latin Extended-F

10780 **107BF**

	1078	1079	107A	107B
0	𐞀	𐞐	𐞠	𐞰
1	𐞁	𐞑	𐞡	
2	𐞂	𐞒	𐞢	𐞲
3	𐞃	𐞓	𐞣	𐞳
4	𐞄	𐞔	𐞤	𐞴
5	𐞅	𐞕	𐞥	𐞵
6		𐞖	𐞦	𐞶
7	𐞇	𐞗	𐞧	𐞷
8	𐞈	𐞘	𐞨	𐞸
9	𐞉	𐞙	𐞩	𐞹
A	𐞊	𐞚	𐞪	𐞺
B	𐞋	𐞛	𐞫	ɟ
C	𐞌	𐞜	𐞬	ƹ
D	𐞍	𐞝	𐞭	ʄ
E	𐞎	𐞞	𐞮	c
F	𐞏	𐞟	𐞯	ʞ

Figures

Although only a few modifier click letters have been attested, these figures show productive use, and that the missing letters are accidental gaps.

In their chapter on aberrant speech, Kelly & Local (1989) give transcriptions with superscript click letters. The trsanscriptions are taken from field notes rather than being artificially contrived. Two click letters are seen. One, dental ⟨ᵻ⟩, is requested here. The other, bilabial ⟨ᵒ⟩, was requested in L2/20-253 and is encoded at U+107B5.

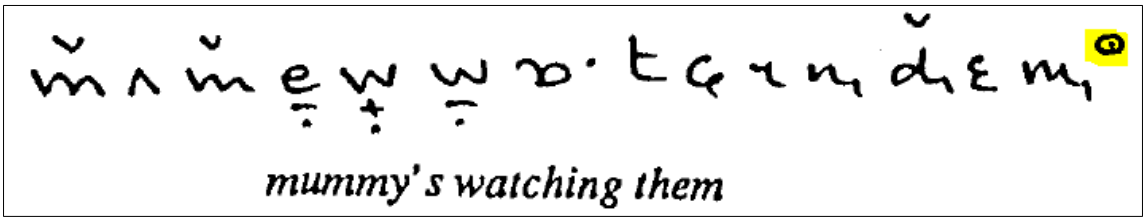

Figure 1. Kelly & Local (1989: 193, 196) ends in a light bilabial click release, U+107B5 ⟨ᵒ⟩, matching the place of articulation of the final nasal consonant.

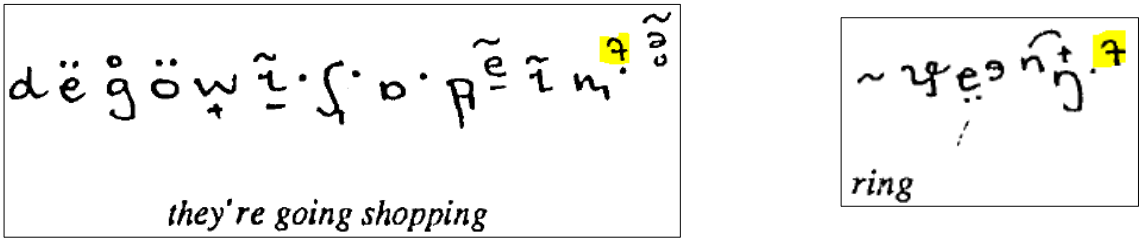

Figure 2. Kelly & Local (1989: 194, 154). The utterances end in a light dental click release ⟨ᵻ⟩, matching the places of articulation of the final consonants.

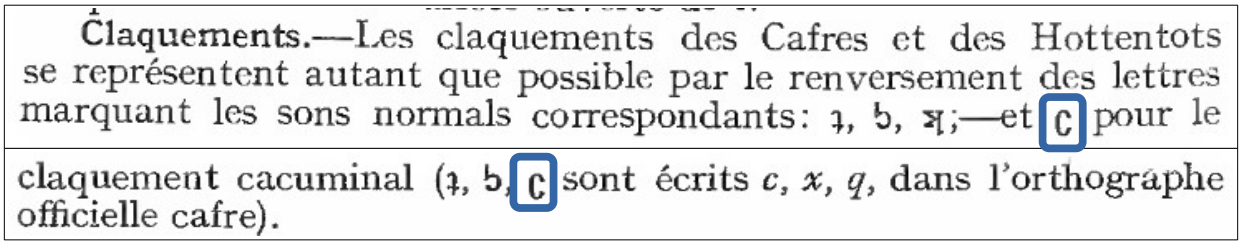
Claquements.—Les claquements des Cafres et des Hottentots se représentent autant que possible par le renversement des lettres marquant les sons normals correspondants: ɬ, ɓ, ǂ;—et ᴄ pour le claquement cacuminal (ɬ, ɓ, ᴄ sont écrits *c*, *x*, *q*, dans l'orthographe officielle cafre).

Figure 3. Passy & Jones (1921: 8–9). The original proposal for the pre-Kiel click letters. ⟨ᴄ⟩ is slightly rounded and full width, very similar in shape to the glyph in Gentium font.

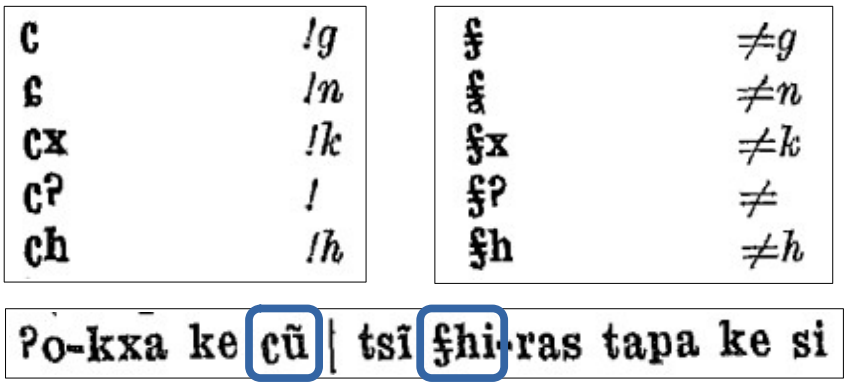

Figure 4. Beach (1938: xiv–xv). Beach used a slightly rounded stretched *c* and a very wavy double-barred esh; the latter is his invention, and matches his esh in shape.

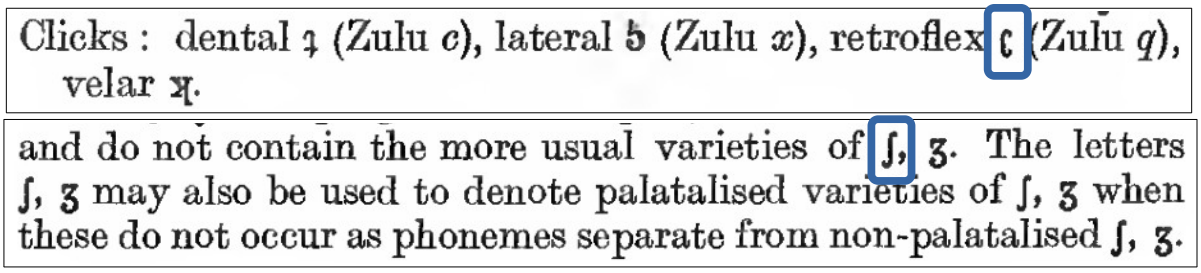

Figure 5. IPA (1949: 13–14). ⟨ɕ⟩ and ⟨ʃ⟩ in IPA publications have a straight vertical stretch and are compressed laterally compared to unstretched ⟨c⟩ and ⟨s⟩.

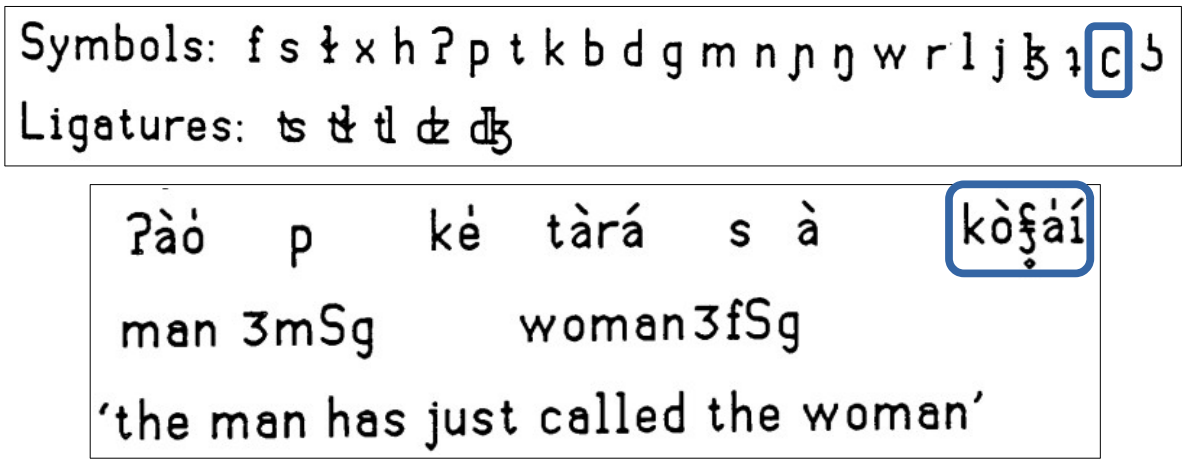

Figure 6. Elderkin (1989: 23, 216). Late use of the pre-Kiel click letters. The ⟨ɕ⟩ has a straight vertical but is not compressed horizontally, while the ⟨ʂ⟩ is as wavy as the glyph in Beach (1938).

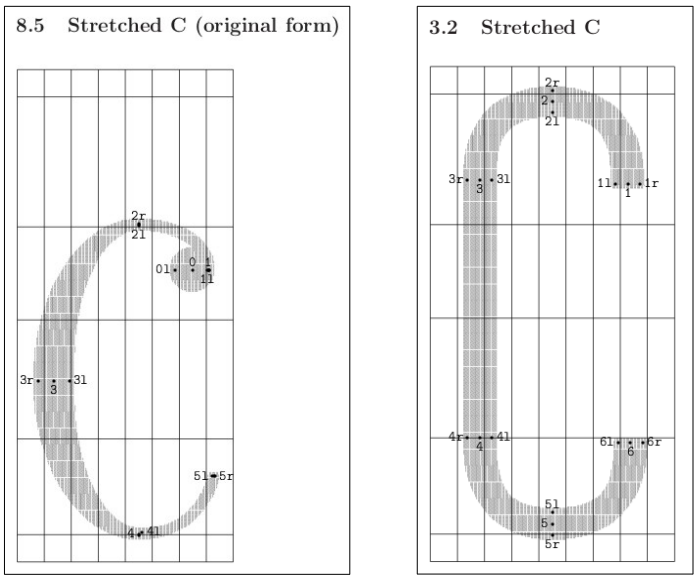

Figure 7. Fukui (n.d. 134, 346). The curved ('original') and straight forms of the letter ⟨ɕ⟩ in TIPA typeface. The ascender on the straight form is not found in pre-Kiel IPA publications, which show only a descender.

ISO/IEC JTC 1/SC 2/WG 2
PROPOSAL SUMMARY FORM TO ACCOMPANY SUBMISSIONS
FOR ADDITIONS TO THE REPERTOIRE OF ISO/IEC 10646[1]
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:** *Modifier pre-Kiel click letters*
2. Requester's name: *Kirk Miller*
3. Requester type (Member body/Liaison/Individual contribution): *individual*
4. Submission date: *2024 February 21*
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: *Latin Extended-F*
2. Number of characters in proposal: *5*
3. Proposed category (select one from below - see section 2.2 of P&P document):
   A-Contemporary x 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)? *yes*

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? *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? *The International Phonetic Organization*
   If YES, available relevant documents: *(see letter of support)*
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) *phonetic*
   Reference:
5. Are the proposed characters in current use by the user community? *yes*
   If YES, where? Reference: *see illustrations*
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)? *partially*
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: