

Unicode request for additional phonetic click letters

Kirk Miller, kirkmiller@gmail.com

Bonny Sands, sands@lowell.edu

2020 April 14

This request is for phonetic symbols used for click consonants, primarily for Khoisan and Bantu languages. For the two symbols <ɘ> and <ɤ>, which fit the pre- and post-Kiel IPA and see relatively frequent use, code points in the Basic Plane are requested. Per the recommendation of the Script Ad Hoc Committee, two other symbols, <!!> and <|||>, may be best handled by adding annotations to existing punctuation and mathematical characters. The remainder of the symbols being proposed are best placed in a supplemental plane. Three precomposed characters, old IPA ɘ ɤ ɥ with a swash (ɘ ɤ ɥ), are covered in a separate proposal.

Characters

- | | |
|---|--|
| ƒ | U+A7F0 LATIN SMALL LETTER ESH WITH DOUBLE BAR. Figures 1–9. |
| ƒ | U+10780 LATIN SMALL LETTER ESH WITH DOUBLE BAR AND CURL. Figures 8–10. |
| ɹ | U+10781 LATIN SMALL LETTER TURNED T WITH CURL. Figures 8–13. |
| Ĉ | U+10782 LATIN LETTER STRETCHED C WITH CURL. Figures 8, 9, 12, 14. |
| ɖ | U+10783 LATIN LETTER INVERTED GLOTTAL STOP WITH CURL. Figures 8, 9, 11, 12, 15. |
| ɭ | U+A7F1 LATIN LETTER RETROFLEX CLICK WITH RETROFLEX HOOK. Figures 16–19. |
| Ɑ | U+ sup plane so U can join it . LATIN LETTER SMALL CAPITAL TURNED K. Figure 20. |

Suggested annotations

U+A7F0 LATIN SMALL LETTER ESH WITH DOUBLE BAR ⟨ƒ⟩ resembles U+2A0E INTEGRAL WITH DOUBLE STROKE ⟨ƒ⟩, though the glyph design is different. (The Ad Hoc Committee recommended not unifying them.)

U+203C DOUBLE EXCLAMATION MARK <!!> may be used as a phonetic symbol, equivalent to a doubled U+1C3 LATIN LETTER RETROFLEX CLICK <!>. A single character is required when modifying by a tie bar, U+0361 COMBINING DOUBLE INVERTED BREVE or U+035C COMBINING DOUBLE BREVE BELOW, as well as for voicing and nasalization diacritics.

U+2980 TRIPLE VERTICAL BAR OPERATOR (≡) may be used as a phonetic symbol, forming a graphic set with U+1C0 LATIN LETTER DENTAL CLICK (ǀ) and U+1C1 LATIN LETTER LATERAL CLICK (ǁ). A single character is required when modifying by a tie bar, U+0361 COMBINING DOUBLE INVERTED BREVE or U+035C COMBINING DOUBLE BREVE BELOW, as well as for voicing and nasalization diacritics.

(This last might be considered for U+2AF4 TRIPLE VERTICAL BAR BINARY RELATION $\langle\equiv\rangle$ or U+2AFC LARGE TRIPLE VERTICAL BAR DELIMITER $\langle\equiv\equiv\equiv\rangle$. Whether the existing clicks better match the height of U+2980 or U+2AFC depends on the font – originally, the click letters had no descender, and so better matched U+2980, but Gentium gives them descenders, better matching U+2AFC.)

Properties

A7F0;LATIN SMALL LETTER ESH WITH DOUBLE BAR;Ll;0;L;;;;N;;;;;
10780;LATIN SMALL LETTER ESH WITH DOUBLE BAR AND
CURL;Ll;0;L;;;;N;;;;;
U+10781;LATIN SMALL LETTER TURNED T WITH CURL;Ll;0;L;;;;N;;;;;
U+10782;LATIN LETTER STRETCHED C WITH CURL;Ll;0;L;;;;N;;;;;
U+10783;LATIN LETTER INVERTED GLOTTAL STOP WITH
CURL;Ll;0;L;;;;N;;;;;
U+A7F1;LATIN LETTER RETROFLEX CLICK WITH RETROFLEX
HOOK;Ll;0;L;;;;N;;;;;
????;LATIN LETTER SMALL CAPITAL TURNED K;Ll;0;L;;;;N;;;;;

Figures

Esh with double bar (Ƀ)

Beach proposed this as a graphically consistent variant of Ƀ to be used in the IPA alongside Ƀ Ƀ. It has been used by other linguists for palatal clicks in texts that use ex-IPA Ƀ Ƀ. John Wells (p.c. 2015 oct 29) said, “It would be good to have Beach’s symbol available in Unicode.”

Though semantically the same as Ƀ, the graphic distinction between Ƀ and Ƀ is equivalent to that of Ƀ Ƀ and Ƀ Ƀ. The letter is graphically similar to the integral sign U+2A0E <Ƀ>, but the Script Ad Hoc Committee decided that they were graphically distinct, and that the integral sign is not an appropriate substitute for the click letter.

A proper <Ƀ> would be useful beyond current use. Amanda Miller (no relation, p.c. 2012) said,

I like the idea of using Beach’s 1938 palatal click symbol for the fricated post-alveolar click type found in Ekoka !Xun, and I would appreciate it if you would get the SIL to include that symbol in the new fonts that they are developing. That would save me a lot of work!

If <Ƀ> were added to Unicode, we’d have all of Beach’s oral click letters. The nasal letters (formed with a loop at the bottom, see next entry) would still not be supported, but there is precedent of using the oral Beach click letters without the nasal letters. Elderkin (1983) *Tanzanian and Ugandan Isolates*, for example, uses a preceding nasal superscript, <Ƀⁿ Ƀⁿ Ƀⁿ Ƀ>, for nasal clicks. He continued to use “the symbols of the prelapsarian IPA” after the IPA changed over, retaining Ƀ alongside Ƀ Ƀ despite not using the dedicated letters for nasal clicks in that publication.

koa (praise), Ƀoa (threaten), Ƀoa (relate), kɃoa-kɃoa (refresh) ɃɃɃɃ (elephant), tɃoa-tɃoa (begin).

Figure 1. Beach (1938: 51). Double-barred esh alongside former IPA letters Ƀ and Ƀ.

250 THE PHONETICS OF THE HOTTENTOT LANGUAGE

(3) ɃN.
ɃɃi (blind), Ƀai (strong), Ƀuru (thunder).

(4) ɃN.
Ƀans (chin), Ƀora (divide), Ƀupu (knock).

Figure 2. Beach (1938: 250)

Class 2 root.		Class 5 root.	
1.	ɃɃa (burst)	ɃɃapa (burst)	
2.	ɃɃa (slaughter)	ɃɃapa (cut open)	
3.	Ƀha (broad)	Ƀhapa (broad)	
4.	ɃɃa (sharp)	ɃɃara (vexatious)	

Figure 3. Beach (1938: 276)

Nasal-curl click letters (ǀ ǃ ǂ Ǆ)

These complete the set of eight click letters created by Beach. Other linguists used them for at least half a century, both when citing Beach for Khoekhoe and when publishing original fieldwork of other languages.

(a) Oral finals.				
(1)	ai,	as in	ǂhai	(run away)
(2)	ae,	„ „	ǂae	(sing)
(3)	ao,	„ „	ǂao	(cling)
(4)	au,	„ „	ǂau	(strike)
(5)	oe,	„ „	ǂ?oe	(remain out)
(6)	oa,	„ „	ǂhoa	(crooked)
(7)	ui,	„ „	ǂhui	(explode)
(b) Nasal finals.				
(8)	āi,	as in	ǂ?āi	(think)
(9)	āũ,	„ „	ǂāũp	(hill)
(10)	ōā,	„ „	ǂhōās	(message)
(11)	ūi,	„ „	ǂūis	(spider).

Figure 8. Beach (1938: 261), phonemic inventory. Also illustrates plain <ǂ>. (The lack of serif at the top of the <ǂ> is just a detail of the font; the letter derives from <ǂ>.)

THE PHONETIC HISTORY OF THE HOTTENTOT ROOTS 261

(2) Click initials.

In Nama (20):

ǂ, ǂ, ǂx, ǂh, ǂ?, ǂ, ǂ, ǂx, ǂh, ǂ?, ǂ, ǂ, ǂx, ǂh, ǂ?, c, ǂ, ǂx, ǂh, ǂ?

In Korana (20 or 24)

Figure 9. Beach (1938: 266), showing all four nasal click letters in transcription.

out his contention, others require rather too great a stretch of the imagination. His examples are ǂam (whistle), ǂom (laugh), sam (skin, bark), ǂxam (hold fast), ǂom (hairy), ǂham (pour). There are many other roots ending in m which could

Figure 10. Beach (1938: 255), with <ǂǂ>.

There is one item which shows clearly how a diphthong functions as a vowel in that it can cooccur with nasalisation. It is ǀǂǂǂǂǂǂ 'the other side of'. ǀǂǂǂ does not occur outside this context although ǂǂǂǂ is a common locative suffix, (to be precise, a sequence of two locative suffixes). Compare Nama ǂǂǂǂ 'in' and ǂǂǂǂ/hǂǂǂǂ nǂǂǂǂ 'within'. I also noted the Sandawe form as ǂǂǂǂǂǂǂǂ, but I do not know if I was accurate, or if indeed, the form was an attempt to adapt ǂǂǂǂǂǂǂǂ to the permitted phonetic syllable structure by reanalysing it as ǂǂǂǂǂǂǂǂ, (from ǂǂǂǂǂǂǂǂ+ǂǂǂǂǂǂǂ), which would be a possible sequence. This tonal pattern might also have been because of the following ǂǂ, (irregularly). ǂǂǂǂǂǂǂǂ is the only item I have found with a nasalised diphthong.

Figure 14. Elderkin (1989: 211), with <ǂ>.

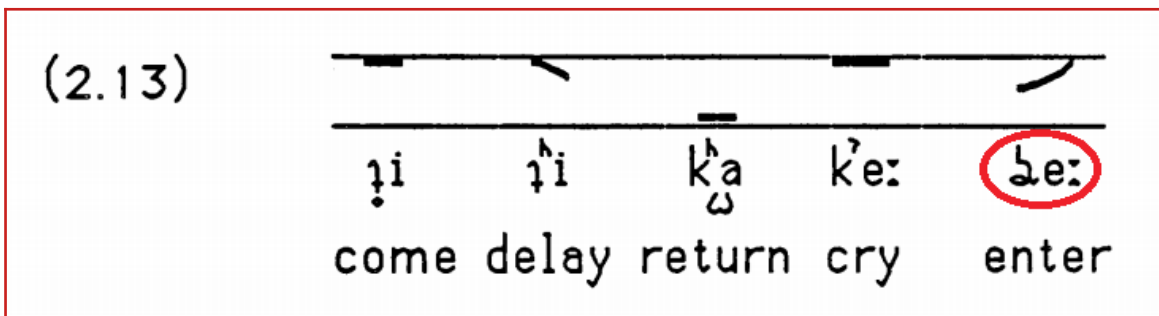


Figure 15. Elderkin (1989: 42), with <ǂ>.

! with retroflex tail (ǀ)

Attested from !Kung. Would be 'sanctioned IPA transcription' (Sands forthcoming, Figure 19).

ǀ A sixth click, which has not received an IPA symbol but is sometimes notated [!] or [!] is the true retroflex click. This is similar to [!], but the tongue tip is placed a little further back, and the contact may be apical or sublaminal. The impression is slightly softer and higher than [!], and in the Khoisan languages and dialects in which it appears, it corresponds to [ʘ] in the

Figure 16. Bradfield (2014: 3)

Table 36.1. Click symbols

	Bilabial	Dental	Alveolar	Retroflex	Palato-alveolar	
voiceless	ɔ		!	ɿ	†	
voiced	ɡɔ	ɡ	ɡ!	ɡɿ	ɡ†	
voiced	ŋɔ	ŋ	ŋ!	ŋɿ	ŋ†	nasal
voiceless						lateral
						click

Figure 17. Bickford & Floyd (2006: 186).

Bilabial	Dental	Alveolar	Alveolar lateral	Retroflex	Palato-alveolar
ŋɔa	ŋ a	ŋ!a	ŋ a	ŋɿa	ŋ†a
aŋɔa	aŋ a	aŋ!a	aŋ a	aŋɿa	aŋ†a
ɡɔa	ɡ a	ɡ!a	ɡ a	ɡɿa	ɡ†a
aɡɔa	aɡ a	aɡ!a	aɡ a	aɡɿa	aɡ†a

Figure 18. Bickford & Floyd (2006: 187)

and Heine 2008: 5). A retroflex hook on the dot on the symbol for a (post)alveolar click  would be a sanctioned IPA transcription but it is not as distinctive looking as .

Figure 19. Sands (forthcoming)

Small capital turned K (x)

As noted in Heselwood (2013: 122), turned κ was mentioned in the 1949 IPA *Principles* (p. 19) as a suggestion for a generic consonant symbol, alongside <ɒ> for a generic vowel symbol, now usually C and V. More recently, <x> has been used as a generic symbol for clicks that allows one to define an ‘accompaniment’ without specifying a place of articulation.

phonological as well as phonetic clusters. Here I just describe the phonetics.

The **Ɂ** fricative clicks in rows 22–23 are so notated because the fricative is fairly long and prominent, making **Ɂ** more descriptive than the possible alternative **Ɂ̥** suggesting an affricated posterior release. As I discuss below, there are also systematic reasons for treating them as a click followed by a fricative.

The **Ɂh** clicks in rows 24–25 have received special attention in the phonetic literature. This, or a similar **Ɂh** accompaniment is found in other languages, including Khoekhoe. It has a distinctive auditory impression, as one hears a long crescendo aspiration (some 200 ms, sometimes even 400 ms) after the click; but the posterior release is not audible. For Khoekhoe (Nama), Ladefoged and Traill (1984) used airflow measurements to establish that the silent start is achieved by nasal venting during the click **Ɂ̥h** for !Xóõ, Traill (1991) showed that this is supplemented by breathing in during the click (so **Ɂ̥↓h**), making it the only established example of ingressive pulmonic airflow in normal language.² There is a question about whether the nasalization is phonetic or phonological, which will be touched on below. I treat it as phonetic, and do not write it.

The clicks **Ɂʔ** with glottal stop in rows 26–27 also tend to have nasalization, at least in the voiced version, and this may or may not be phonological – here I have assumed not. They are auditorily distinguished from the ejectives **Ɂ'** in rows 5–6 mainly by the lack of an audible posterior release – similar to the difference between saying [ak'a] and [akʔa].

Figure 20. Bradfield (2014: 9). Small-capital turned K representing various families of clicks.

For future reference

Old Zulu click letters (Ɂ, Ɂ̥)

Used by the 19th-century Norwegian mission to the Zulu. No known interest in digitizing such material at present. The first letter resembles a Greek qoppa, Ɂ, in some fonts, but qoppa has the wrong height and will not have the proper shape in all fonts.

	Udtale		Udtale		Udtale		Udtale		Udtale		Udtale
a	a	g	ge	k	ke	Ɂ	thle	Ɂ	rhe	w	we
b	be	h	he	Ɂ̥	kje	m	em	s	se	Ɂ̥	ve
d	de	i	i	Ɂ̥	kje	n	en	Ɂ̥	se	z	ze
e	e	j	je	l	le	o	o	t	te		
f	fe	Ɂ̥	dhje	Ɂ̥	dhle	p	pe	u	u		
samt de uartikulerte Lyd. Ɂ. Ɂ̥. Ɂ̥̥.											

Figure 21. Schreuder & Holmboe (1850: 1). The barred letters of the alphabet are covered by Unicode *j k l*, with U+A7CA planned for s-bar. The diacritics might be rendered as *k̥ k̥* etc., though the typesetter here apparently used overscript *k̥ k̥*.




ere eiendommelige smekkende, klikkende og klukkende
Lyd;  udtalt *foran* (dentale),  *ved Siden* (laterale) og 
bag i Munden (guttural-palatin); desuden knytter der sig til

Figure 22. Schreuder & Holmboe (1850: 5). Single <ɘ> is modern c; double <ɘɘ> (perhaps inspired by Greek ξ?) is modern x. <ɘ̣> for modern q is distinguished by a diacritic and so should not need separate Unicode encoding.

— 25 —




kafula plur.; Iɘosa, Amaɘosa Navnet paa en Kafferstamme;
Izulu, Amazulu, Amazuluer; amasole date Soldater etc.

Figure 23. Schreuder & Holmboe (1850: 25). <Iɘosa> *Ixosa* and <Amaɘosa> *Amaxosa* (“Xhosa”) in the Norwegian Zulu alphabet, identifying <ɘ> with modern x.

Gamaɘebo ake wa ba kofisa ukw apula izimemezelo zi ka Jehova.
Ved Knebene sine han dem bedrog at bryde Budene dem Guds.
Abantu bokuɘala ba nga lindanga izinɘizijo zabo, ngako ba ɘukwa
Menneskene de første de ikke bevogtede Hjerterne sine, derfor de fangedes
gamaɘebo ka Satane, ba z' apula izimemezelo zenkosi, ba wa

Figure 24. Schreuder & Holmboe (1850: 83). Textual use of the click letters, with <gamaɘebo> and <ɘukwa> for modern c and <bokuɘala> for modern q.

References

- Douglas Beach (1938) *The Phonetics of the Hottentot Language*.
William Bennet (forthcoming) ‘Click phonology’. *Click Consonants*, Brill, Leiden.
Bickford & Floyd (2006) *Articulatory Phonetics*, 4th ed.
Julian Bradfield (2014) “Clicks, Concurrency and Khoisan,” *Phonology*, vol. 31, no. 1.
Derek Elderkin (1989) *The Significance and Origin of the Use of Pitch in Sandawe*.
Sands (forthcoming) ‘Introduction’. *Click Consonants*, Brill, Leiden.
Schreuder & Holmboe (1850) *Grammatik for Zulu-Sproget*, Christiania.

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 <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:	Additional phonetic click letters	
2. Requester's name:	Kirk Miller, Bonny Sands	
3. Requester type (Member body/Liaison/Individual contribution):	individual	
4. Submission date:	2020 March 24	
5. Requester's reference (if applicable):		
6. Choose one of the following:		
This is a complete proposal:	x	
(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):		
Proposed name of script:		
b. The proposal is for addition of character(s) to an existing block:	x	
Name of the existing block:	Latin Extended-D, Phonetic Extensions Supplement-A	
2. Number of characters in proposal:	7	
3. Proposed category (select one from below - see section 2.2 of P&P document):		
A-Contemporary	x	B.1-Specialized (small collection)
C-Major extinct		B.2-Specialized (large collection)
F-Archaic Hieroglyphic or Ideographic		E-Minor extinct
		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 <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.

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?	<i>The authors are a members of the user community.</i>
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)	<i>phonetic</i>
Reference:	
5. Are the proposed characters in current use by the user community?	yes
If YES, where? Reference:	<i>see illustrations</i>
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:	<i>(annotations to support those that are)</i>
9. Can any of the proposed characters be encoded using a composed character sequence of either existing characters or other proposed characters?	yes
If YES, is a rationale for its inclusion provided?	
If YES, reference:	<i>(Unicode disprefers use of combining retroflex tail)</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?	yes
If YES, is a rationale for its inclusion provided?	yes
If YES, reference:	<i>(see comment by Script Ad Hoc Committee on integral sign)</i>
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:	