

Unicode request for extIPA support

Kirk Miller, kirkmiller@gmail.com

Martin Ball, m.j.ball@bangor.ac.uk

2020 Jan 08

This is a request for additions to Unicode to support recent expansion of the extIPA (i.e., Extensions to the IPA for Disordered Speech), including superscript modifier letters for unusual releases of plosives. The full 2018 JIPA article on the 2015 revision of the extIPA is publicly available online from Cambridge University Press. (See references.)

Thanks to Deborah Anderson of the Universal Scripts Project for her assistance.

As summarized in the JIPA article, *the motivation behind the original 1990 version of the extIPA symbols, and subsequent revisions to them, has been to supply transcriptional resources to those needing to describe disordered speech of all types. Symbols have been chosen following requests from speech-language pathologists and clinical phoneticians.*

As for the recent update, *at the Oslo conference of the International Clinical Phonetics and Linguistics Association (ICPLA) in 2010, a special panel on the transcription of disordered speech reviewed the extIPA symbols and the chart that displays them, with input also from the audience at the panel presentation (Ball et al. 2010). The results of this review, together with informal consultations with colleagues, led to a set of suggested changes. These were presented in a poster in June 2016 at the ICPLA conference in Halifax, Nova Scotia (Ball et al. 2016), and approved by the membership. The changes were presented as a motion at the Business Meeting (all ICPLA members present at conference) and approved nem. con.*

Inline letters

- ɔ turned small capital G
- ɹ reversed k
- ɔ̣ reversed script g
- ɹ̣ reversed eng
- ff̣ f-eng ligature with trill
- ḷ l-ezh ligature with right tail
- ɣ turned y with belt
- ɬ small capital L with belt

All are shown in Figures 1 and 3. In addition, use of ɣ in non-disordered speech is shown in Figures 4–6, and use of ɬ in Figures 7–9.

Modifier letters

- ɬ modifier l with belt
- ḷ modifier l-ezh ligature
- ɬ̣ modifier l with belt and right tail
- ḷ̣ modifier l-ezh ligature with right tail

- ɣ modifier turned y with belt
- ɛ modifier small capital L with belt
- ɬ modifier l-s ligature
- ɺ modifier l-z ligature

Of these, ɛ, ɬ, ɺ are shown in Figure 2, while the others are covered by the provision, in the JIPA article and chart, for superscript lateral fricatives to indicate ‘lateral fricated release’ of a plosive (article quoted under ‘Superscript lateral fricatives’ below, after Figure 9).

ɧ is shown in Figure 10 *ff*, which also show cases where ɧ and ɣ would have been used if the inline characters had been available at the time. ɬ is shown in Figure 11.

Question for *ad hoc* committee

Given that there isn’t space in the Basic Plane for IPA superscripts, let alone for subscripts or to handle Americanist, Danish and other phonetic traditions, might it be better to use something like **variation selectors** so that super-/sub-/etc. scripting could be assigned to any character?

In addition to 86 baseline phonetic characters being proposed by Miller, so far there are 64 superscripts (not counting ɧ ɬ ɣ above or letters like ɔ ɔ̥ ɔ̥̥ that are potentially composite) and 21 subscripts, and we’re continually finding more. Creating variation selectors for small and petite caps would handle a dozen of the Americanist symbols as well, e.g. for voiceless vowels. One of the unassigned 16-point Basic Plane blocks should be able to handle all of these, as well as the italic, double-struck, cursive, Fraktur etc. variants that are often needed for technical symbols.

Or perhaps for modifier phonetic letters and non-IPA traditions, we need to open blocks in one of the extended plains. But variation selectors would seem to be a more efficient approach. And a good word-processor could handle them with soft formatting even if a particular font does not handle them directly.

Combining diacritics

- Ⓟ combining subscript left parenthesis
- Ⓠ combining subscript right parenthesis
- Ⓡ combining left parenthesis
- Ⓢ combining right parenthesis

These unpaired equivalents of existing diacritics are shown in Figure 3.

Figures

ExtIPA charts

CONSONANTS (other than those on the IPA Chart)

	Bilabial	Labio-dental	Labio-alveolar	Dento-labial	Bidental	Linguo-labial	Inter-dental	Alveolar	Retroflex	Palatal	Velar	Velo-pharyngeal	(Upper) pharyngeal
Plosive		p̣ ḅ	p̥ b̥	p̄ b̄		ṭ ḍ	t̥ d̥						q̣ ɢ
Nasal			ṃ ṃ	m̥ m̥		ṇ ṇ	n̥ n̥						
Trill						ṛ	r̥					ʀ̥ ʀ̥	
Fricative, median			f̣ ṿ	f̥ v̥	ɦ̣ ɦ̣	θ̣ ð̣	θ̥ ð̥	θ̄ ð̄				ɸ̣ ɸ̣	
Fricative, lateral						ɬ̣ ɬ̣	ɬ̥ ɬ̥		ɬ̣ ɬ̣	ʃ̣ ʃ̣	ʂ̣ ʂ̣		
Fricative, lat. + med.								ḷ ṣ					
Fricative, nasal	ṃ ṃ	ṃ ṃ						ṇ ṇ	ṇ ṇ	ɲ̣ ɲ̣	ŋ̣ ŋ̣		
Approx., lateral						ḷ	ḷ						
Percussive	ẉ				ɸ̣								

Figure 1. Ball et al. (2018), p. 160, main consonant chart. Unsupported characters are <ɢ> and <ɬ̣, ʃ̣, ʂ̣> (described on p. 161, sect. 3.2 para 1), and <ʀ̥> (described on p. 161, sect. 3.2 para 3).

OTHER SOUNDS

ɹ̣	apical-r	ṭθ̣	linguolabial affricate etc.
ɹ̣	bunched-r (molar-r)	ɬ̣ ɬ̣ ɬ̣	velodorsal oral and nasal stops
ɸ̣ ɸ̣	laminal fricatives (incl. lowered tongue tip)	ɹ̣	sublaminal lower alveolar percussi
ḳ ^{ɬ̣} etc.	[k] with lateral fricated release etc.	!ɹ̣	alveolar click with sublaminal perc
ṭ ^{ḷ} ḍ ^{ḷ}	[t, d] with lateral and median release	ʋ̣	buccal interdental trill (raspberry)
ṭ ^{ɦ̣}	[t] with interdental aspiration etc.	*	sound with no available symbol

Figure 2. Ball et al. (2018), p. 160, Other Sounds. The full letters <ɬ̣, ɸ̣, ɹ̣> are not supported. These are described on the top of p. 161. In addition, modifier variants of the new laterals are illustrated as consonant releases. Although both lateral+medial fricatives are given in the chart, as <ṭ^{ḷ}, ḍ^{ḷ}>, only a sample of the plain laterals is given, <ḳ^{ɬ̣}>, with the word “etc.” added to show this is a general pattern.

Ejectives (cf. 2), fricatives (cf. 3) and prenasalized stops (cf. 2), each only contrast at three places of articulation, while affricates (cf. 3) contrast at only two places of articulation.

(3)	Labio-dental	Alveolar	Palato-alveolar	
Prenasalized affricate		nts ndz	ndʒ	
Central affricate		ts dz	dʒ	
Lateral affricate			tʃ	= tʃ (tʃ _l)
Ejective central affricate		ts'	tʃ'	
Ejective lateral affricate			tʃ'	= tʃ'
Central fricative	f	s	ʃ	
Lateral fricative		ɬ		

Hadza has a three-way Voice Onset Time (VOT) contrast in the plosives, i.e. voiceless

Figure 4. Sands, Bonny, 2013. *Phonetics and Phonology: Hadza*. In Rainer Vossen (ed.) *The Khoesan Languages*. Routledge, p. 39. The palatal lateral fricative <ʃ> used as a component of the corresponding affricates. The barely legible form created by the publisher was clarified in a margin note by the author before she distributed copies to colleagues.

1.3.2 Vocalic variation

Vowel height harmony takes place across morpheme boundaries within a phonological word. Mid vowels /e o/ are raised to [i] and [u] before a high vowel /i u/, e.g. /hanj/ 'akwe + p^hi/ [hanj^h'akwip^hi] 'are gullets', /hik'o + p^hi/ [hik^h'op^hi] 'are wooden-tipped arrows'. Vowel harmony can spread to earlier syllables, e.g. /loʔo + p^hi/ [loʔop^hi] 'are long horns', but cf. /tʃ'oma + p^hi/ [tʃ'omap^hi] 'are heads'.

Final vowels often become voiceless [ɿ ɛ ɔ ɔ̥ ʊ] in final position, particularly when preceded by a glottal stop or any other voiceless stop.

Figure 5. Sands (2013), p. 42.

dlaa [cʃ'aʔa] [verb] to sing. *dlaâmo*. *Dlaate onebee* 'let us sing' (IMP.2pl for HORT.1incl)

dladlaa, -ko (dla~dlaa) [cʃ'acʃ'a'ʔa] [not a long a] a singer (s.o. good at singing, whether or not they're known, or sing often.)

dladlaangu (dla~dlaa-ngu) [cʃ'acʃ'a'ʔaŋɡu] an ever better singer. positive connotation.

dladlafengu (dla~dlaa-fe-ngu) [cʃ'acʃ'a'ʔa'fɛŋɡu] s.o. who only sings, doesn't work or help out. negative connotation.

Figure 6. Miller, Kirk, Mariamu Anyawire, G.G. Bala & Bonny Sands, forthcoming. *A Hadza Lexicon and Etymological Dictionary*. First entry under dictionary heading **DL**, from the ms distributed in 2016.

Attempts to use the letter <ɬ> in published descriptions of Hadza have so far failed due to a lack of Unicode support, though this will hopefully change with the forthcoming dictionary.

1.3.1 Consonantal variation

The alveolar click /!/ can be produced with a range of phonetic realizations, from a very quiet, weak click, to a loud pop, or as a noisy “flapped” or “plopped” click in which the tongue tip makes contact with the bottom of the mouth after the release of the front click closure. Denti-alveolar, alveolar, and post-alveolar closures for /!/ are in free variation.

In intervocalic position, the approximant [l] can appear as [r].

Ejective velar /k'/ is pronounced with heavy frication [kx' kɬ'] by some speakers.

The voiceless nasalized clicks /ŋ!' ŋ!' ŋ!' are always produced with voiceless nasal airflow during the click closure. (not for all speakers, apparently)

When a vowel precedes the click, either within a morpheme, e.g. [hãŋ!'a-k^ho] ‘rock’, or across morpheme boundaries, the vowel becomes partially nasalized.

An initial aspirated obstruent or voiceless click will have a shorter VOT (i.e. less aspiration) than the other clicks.

kɬ',
not
kx'

Figure 7. Sands (2013), p. 41: The velar letter <ɬ> was mistakenly replaced by palatal <ɧ> by the publisher. The submitted manuscript had used <ɬ> (see next figure). The author manually corrected the letter in a margin note before distributing copies to her colleagues.

III. 1. Consonantal variation

- The alveolar click /!/ can be produced with a range of phonetic realizations, from a very quiet, weak click, to a loud pop, or as a noisy “flapped” or “plopped” click in which the tongue tip makes contact with the bottom of the mouth after the release of the front click closure. Denti-alveolar, alveolar, and post-alveolar closures for /!/ are in free variation.
- In intervocalic position, the approximant [l] can appear as [r].
- Ejective velar /k'/ is pronounced with heavy frication [kx' kɬ'] by some speakers.
- The voiceless nasalized clicks, /ŋ!' ŋ!' ŋ!' are always produced with voiceless nasal airflow during the click closure, but when a vowel precedes the click, either within a morpheme, e.g. [hãŋ!'a-k^ho] ‘rock’, or across morpheme boundaries, the vowel becomes partially nasalized.

Figure 8. Sands (n.d.), manuscript of previous figure, showing the intended velar lateral affricate.

gg (k', kx')

An ejective velar stop [k'], affricate [kx'] (sometimes lateral: [kɬ']), or fricative [x']. Typically written <k'> or <kx'> in the anthropological literature. May glottalize adjacent vowels (eg. *beggau* [bek'xʔau] in 2007April09.3, with two clear releases). In Sandawe, [kɬ'] is instead an allophone of *tʰ* before *u*, *w*, but this has not been noticed in Hadza (though see comment at *xxudle*).

Figure 9. Miller et al. (forthcoming). Explanation of dictionary heading **Gg**.

Superscript lateral fricatives (ɬ ɬ̣ ɬ̥ ɬ̧ ɬ̨)

Modifier extIPA letters are described in Ball et al. (2018), p. 162, sect. 3.5 para 2:

The next four lines [of the 'other sounds' chart] illustrate how existing extIPA symbols can be used to transcribe a range of unusual plosive release types: lateral fricated release, lateral and median release, interdental aspiration, and linguolabial affricates.

'Lateral fricated release' does not refer to any particular character or articulation, but to the Lateral Fricative row of the main consonant chart (Fig. 1). In addition to the <k^ɬ> given as an example in the 'other sounds' chart, these include palatal <c^ɬ>, retroflex <ɖ̡, ɖ̢̡>, and alveolar <t^ɬ, d^ɬ> combined with linguolabial and interdental diacritics (the diacritics do not join properly in the current font). [\[add to font\]](#)

These superscript variants are occasionally found outside the extIPA.

Many Mon-Khmer, Berber and Salish languages possess extremely unusual consonant clusters as, for instance, initials in Khasi [bt, dk^h] (Schmidt 1904, Rabel 1961, Henderson 1976a, b), Semai [gpgh, t^stt^sʔ, kdkrld] (Diffloth 1976a, b), Kammu [t^{sh}kb] (Svantesson 1983), Moses-Columbia Salish [xt^ɬ t^ɬ x^w, k't^s] (Czaykowska-Higgins & Willett 1997), Lillooet [q'ʔ] (van Eijk 1997) or obstruent-only words such

Figure 10. <t^ɬ> in Kehrein & Golston (2004) "A prosodic theory of laryngeal contrasts", *Phonology* 21.3, p. 348.

(40) Correspondences of Proto-Athapaskan obstruents in SLAVEY SLAVE (Rice 1989: 89)³⁸

Proto-Athapaskan	Slave	
	initial	final
*t, *d, *t'	t, d, t'	t
*t ^ɬ , *d ^ɬ , *t ^ɬ , *ɬ, *ɬ̣	t ^ɬ , d ^ɬ , t ^ɬ , ɬ	ɬ
*t ^s , *d ^z , *t ^s , *s, *z	t ^θ , d ^θ , t ^θ , θ	θ

Figure 11. Wolfgang Kehrein (2002) *Phonological Representation and Phonetic Phasing*:

Affricates and Laryngeals, Walter de Gruyter, p. 34, with affricates <tʰ> and <dʰ>.

lateral	-	l̥	l	ḷ	l̤
nasal stop	p ^m	t̥ ⁿ	t ⁿ	ṭ ⁿ	t̤ ⁿ
lateral stop	-	t̥ ^l	t ^l	ṭ ^l	t̤ ^l

Figure 12. Kehrein (2002), p. 20. The laterals of Adynyamathanha, with superscript $\bar{\cdot}$ $\bar{\cdot}$. The intervening retroflex sound would presumably have used superscript retroflex $\bar{\cdot}$ if the inline letter had been available at the time.

- 4 IPA does not contain independent symbols for retroflex, palatal, and velar lateral fricatives. Throughout this book I follow Ladefoged & Maddieson (1996) and use the raising diacritic [̥]: thus, I transcribe fricatives as [l̥, ʎ̥, l̥] and affricates as [t̥^l, c^{ʎ̥}, k^{l̥}]. To avoid too much transcriptional complexity I will not further distinguish between voiceless and voiced fricatives (which Ladefoged & Maddieson do by additionally marking the former with [̥]).

Figure 13. Kehrein (2002), p. 6, explaining why they do not use unitary characters for non-alveolar lateral fricatives, though they do use fricative characters in affricates wherever they can.

stops	p	ɸ	t̥	t	ṭ	t̤	c	k
lateral affr	-	-	t̥ ^l	t ^l	ṭ ^l	t̤ ^l	c ^{ʎ̥}	k ^{l̥}
laterals	-	-	l̥	l	ḷ	l̤	ʎ̥	L
lateral frics	-	-	ʎ̥	ʎ̥	ʎ̣	ʎ̤	ʎ̥	L̥

Figure 14. Kehrein (2002), p. 20. A summary of attested laterals, lateral fricatives and lateral affricates. Superscript $\bar{\cdot}$ $\bar{\cdot}$ are circled in red. Superscript $\bar{\cdot}$ $\bar{\cdot}$ are circled in yellow. (Page 6: the ring diacritics for voicelessness are omitted to simplify the transcription.) In several places the author describes a natural set of lateral affricates [t̥^l, t^l, ṭ^l, t̤^l, c^{ʎ̥}, k^{l̥}]. If the extIPA letters $\bar{\cdot}$ $\bar{\cdot}$ had been available, they could have been used as superscripts here. As it is, superscript $\bar{\cdot}$ is used for the palatal. (The retroflex and velar are already supported by Unicode.)



Figure 15. Kehrein (2002), p. 18. A palatal lateral affricate [c^{ʎ̥}].

(51) Stop/affricate alternations in SAHAPTIN 'consonantal ablaut' (Rigsby & Rude 1996: 672)

k' < q': k'uxsk'uxs 'ankle' q'úχł 'knee' (Warm Springs only)

t^s < t^h: kút^s'k 'a smaller piece' kút^h'k 'a piece, chunk'

Figure 16. Kehrein (2002), p. 39. <t^h> in use.

Table 2. Navajo stops in an IPA transcription.

	labial	alveolar	palato-alveolar
plain	p	t ^h t t'	
laterally released		t ^l t ^l t ^l '	
affricated		ts ^h ts ts'	tʃ ^h tʃ tʃ'

Figure 17. McDonough & Ladefoged (1993). Navajo Stops, *UCLA Working Papers in Phonetics* 84, p. 152. The consonant at the left is phonemically /t^h/, phonetically [t^l].

vl. lateral affric. click
 vl. asp. lateral affr. click
 glottalized vl. lat. aff. cl.
 voiced lat. affric. click

ʃ^h

ʃ^h
 ʃ^h?
 ʃ^h

ʃ^h

Figure 18. Lateral affricated alveolar clicks in the UPSID. Maddieson (1981) "UPSID: Data and Index", *UCLA Working Papers in Phonetics* 53, p. 167.

vl. lat. aff. click
 vl. asp. lat. aff. c.
 gl. vl. lat. aff. cl.
 vd. lat. affr. click
 gl. vd. lat. aff. cl.
 breathy vd. lateral
 affricated click

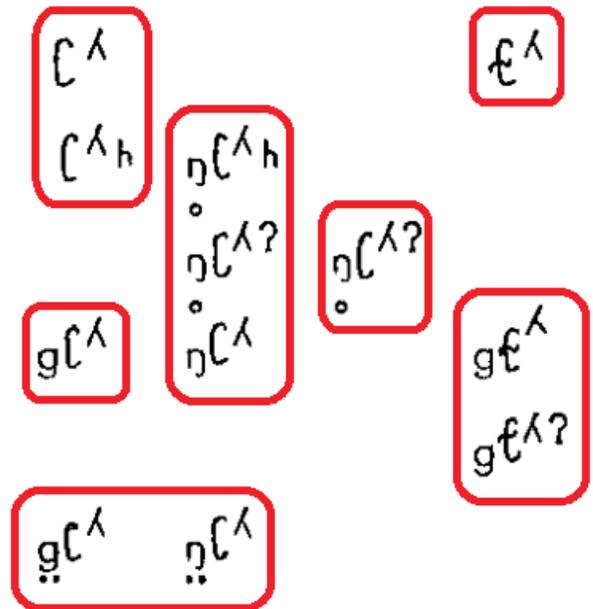


Figure 19. Lateral affricated palatal clicks in the UPSID, p. 170. Frication would more properly be transcribed with a superscript χ . As it is, superscript λ is not found in Unicode either.

References

Ball, Martin J., Sara J. Howard & Kirk Miller, 2018. Revisions to the extIPA chart. *Journal of the International Phonetic Association*, volume 48, issue 2, pp. 155–164, doi: 10.1017/S0025100317000147. Published online by Cambridge University Press, 11 April 2017. <https://www.cambridge.org/core/journals/journal-of-the-international-phonetic-association/article/revisions-to-the-extipa-chart/06C01EA81DA2AECA2AC52AAF21556B33>