Unicode request for letters with overstruck tilde

L2/23-262

Kirk Miller, kirkmiller, gmail.com

2023 October 13

The IPA provides for an overstruck tilde, $\langle * \rangle$, to mark gutturalized (velarized, uvularized and pharyngealized) consonants, prototypically the "dark" (velarized) [\dagger] of English wool but by extension also the "emphatic" (pharyngealized) consonants of Arabic. Semantically, the tilde overlaps with the more precise spacing modifiers $\langle {}^{y}\rangle$ velarized, $\langle {}^{s}\rangle$ uvularized (not included on the IPA chart but listed among the Voice Quality Symbols) and $\langle {}^{c}\rangle$ pharyngealized. An overstruck tilde has occasionally been used in non-IPA traditions for other meanings, such as palatalization.

The tilde itself is encoded at U+0334 COMBINING TILDE OVERLAY. However, proposal N2632 held that dynamic generation of characters with U+0334 should be avoided. Indeed, U+026B LATIN SMALL LETTER L WITH MIDDLE TILDE does not have a decomposition with U+0334, and the precomposed characters presented in proposal N2632 were accepted in Unicode 4.0. These are b df mnprest. A case was made later for a at U+AB68.

Although N2632 included the overstruck IPA letters that the author could identify at the time, a number of others are in use. Some, such as $\langle \eth \rangle$ for the Arabic emphatic consonant \not , are rather common. This proposal covers most cases of unsupported tilde that we could discover, excepting a few *ad hoc* creations (see e.g. Figure 4). Of the basic Latin consonant letters in the IPA, all are now attested with the tilde apart from q and w. ([q] is a uvular consonant – though Figure 44 shows that this is not necessarily an incompatibility – while [w] is inherently velarized.) Among other IPA consonant letters, several voicing pairs are now attested, namely the fricatives articulated at bilabial ($\langle \Phi, B \rangle$), dental ($\langle \Phi, B \rangle$), lateral ($\langle \Phi, B \rangle$) and post-alveolar ($\langle \Phi, B \rangle$) places.

The tilde may also occur on a vowel letter. Woidich (2006a: 26), for example, describes assimilation of the vowel /i/ to the emphatic articulation of neighboring consonants in Arabic: Das /l/ des Artikels assimiliert sich auch an emphatische Konsonanten, das /i/ kann daher emphatisch werden "the /l/ of the article assimilates to an emphatic consonant; the /i/ can therefore become emphatic," and he transcribes that allophone with a tilde: $\langle i \rangle$ (Figure 19). However, when /a/ appears in such a context, Woidich transcribes its emphatic allophone as $\langle a \rangle$ rather than as $\langle a \rangle$ or $\langle a \rangle$. This appears to be a natural gap in the inventory: Catford (1977: 183ff) among others describes the semi-vowel [a] as a pharyngeal approximant, and [a] as a "close" vowel in the sense that the root of the tongue is close to the pharynx. However, Khan (2015) marks low vowels in such contexts with a tilde, so there is evidently some variation in interpretation.

We document three IPA letters with a double overstruck tilde. One exists already as U+AB38 $\frac{1}{4}$, which is the Teuthonista symbol for a very dark $\frac{1}{4}$. It has the same meaning in IPA. The doubling of

an IPA (or Teuthonista) diacritic indicates a greater intensity of that feature. For example, in IPA $\left[\tilde{e}\right]$ is heavily nasal, $\left[a^{\sim}\right]$ is heavily rhotic, $\left[\delta\right]$ has extra-high tone, $\left[\frac{i}{2}\right]$ is highly retracted, etc. In this case, pharyngealized /l/ is treated as a darker sound $\left[\frac{1}{2}\right]$ than velar $\left[\frac{1}{2}\right]$.

Thanks to Denis Moyogo Jacquerye for his feedback and many of the references used below.

Characters

For the placement of the middle tilde, we tried to follow the literature where we could attest to something other than mechanical use of U+0334. Even in manuscript form there is some variability, however, and we leave it an open question whether some of these variants might be disunified in the future if a semantic distinction can be found.

SAH: Should G and SCRIPT G be unified?

- a 1DF3B LATIN SMALL LETTER A WITH MIDDLE TILDE. Figure 20.
- e 1DF3C LATIN SMALL LETTER ALPHA WITH MIDDLE TILDE. Figure 18, Figure 20 ff.
- β 1DF3D LATIN SMALL LETTER BETA WITH MIDDLE TILDE. Figure 1, Figure 39.
- € 1DF3E LATIN CAPITAL LETTER C WITH MIDDLE TILDE. Figure 2.
- ϵ 1DF3F LATIN SMALL LETTER C WITH MIDDLE TILDE. Figure 3.
- 1DF40 LATIN SMALL LETTER STRETCHED C WITH MIDDLE TILDE. Figure 50.
- b 1DF41 LATIN SMALL LETTER INSULAR D WITH MIDDLE TILDE. Figure 12.
- ð 1DF42 LATIN SMALL LETTER ETH WITH MIDDLE TILDE. Figure 4 ff.
- ⇒ 1DF43 LATIN SMALL LETTER SCHWA WITH MIDDLE TILDE. Figure 18, Figure 20 ff.
- g 1DF44 LATIN SMALL LETTER G WITH MIDDLE TILDE. Figure 14 ff.
- g 1DF45 LATIN SMALL LETTER SCRIPT G WITH MIDDLE TILDE. Figure 13, Figure 19.
- † 1DF46 LATIN SMALL LETTER H WITH MIDDLE TILDE. Figure 17 ff, Figure 47.
- i 1DF47 LATIN SMALL LETTER I WITH MIDDLE TILDE. Figure 19.
- 1DF48 LATIN SMALL CAPITAL I WITH MIDDLE TILDE. Figure 19, Figure 24.
- j 1DF49 LATIN SMALL LETTER J WITH MIDDLE TILDE. Figure 24.
- † 1DF4A LATIN SMALL LETTER K WITH MIDDLE TILDE. Figure 15 ff, Figure 19, Figure 25 ff.
- † 1DF4B LATIN SMALL LETTER L WITH PALATAL HOOK AND MIDDLE TILDE. Figure 28.
- † 1DF4C LATIN SMALL LETTER L WITH RETROFLEX HOOK AND MIDDLE TILDE. Figure 29.
- [‡] 1DF4D LATIN SMALL LETTER L WITH BELT AND MIDDLE TILDE. Figure 30 ff.
- 1DF4E LATIN SMALL LETTER LEZH WITH MIDDLE TILDE. Figure 32 ff.
- f 1DF4F LATIN SMALL LETTER TURNED Y WITH MIDDLE TILDE. Figure 45.
- e 1DF50 LATIN SMALL LETTER O WITH MIDDLE TILDE. Figure 20.
- Φ 1DF51 LATIN SMALL LETTER PHI WITH MIDDLE TILDE. Figure 34.
- R 1DF52 LATIN SMALL CAPITAL R WITH MIDDLE TILDE. Figure 35.

- † 1DF53 LATIN SMALL LETTER R WITH TAIL AND MIDDLE TILDE. Figure 36.
- § 1DF54 LATIN SMALL LETTER S WITH HOOK AND MIDDLE TILDE. Figure 37.
- 1DF55 LATIN SMALL LETTER ESH WITH MIDDLE TILDE. Figure 17, Figure 38 ff.
- 1DF56 LATIN SMALL LETTER TURNED T WITH MIDDLE TILDE. Figure 48.
- \$\ 1DF57 LATIN SMALL LETTER TESH DIGRAPH WITH MIDDLE TILDE. Figure 21.
- **♦** 1DF58 LATIN SMALL LETTER THETA WITH MIDDLE TILDE. Figure 11.
- uj 1DF59 LATIN SMALL LETTER TURNED M WITH LONG LEG AND MIDDLE TILDE. Figure 23.
- น 1DF5A LATIN SMALL LETTER U WITH MIDDLE TILDE. Figure 17 ff.
- ช 1DF5B LATIN SMALL LETTER UPSILON WITH MIDDLE TILDE. Figure 19.
- ¥ 1DF5C LATIN SMALL LETTER V WITH MIDDLE TILDE. Figure 18, Figure 20, Figure 30, Figure 40 ff.
- υ 1DF5D LATIN SMALL LETTER V WITH HOOK AND MIDDLE TILDE. Figure 42.
- x 1DF5E LATIN SMALL LETTER X WITH MIDDLE TILDE. Figure 15, Figure 18, Figure 20, Figure 43.
- χ 1DF5F LATIN SMALL LETTER CHI WITH MIDDLE TILDE. Figure 44.
- 3 1DF60 LATIN SMALL LETTER EZH WITH MIDDLE TILDE. Figure 38.
- 2 1DF61 LATIN SMALL LETTER GLOTTAL STOP WITH MIDDLE TILDE. Figure 20, Figure 47.
- 3 1DF62 LATIN SMALL LETTER INVERTED GLOTTAL STOP WITH MIDDLE TILDE. Figure 49.
- 1DF63 LATIN SMALL LETTER N WITH DOUBLE MIDDLE TILDE. Figure 51.
- 1DF64 LATIN SMALL LETTER R WITH DOUBLE MIDDLE TILDE. Figure 51.
- [‡] 1DFFA MODIFIER LETTER SMALL TURNED R WITH MIDDLE TILDE. Figure 37.
- ⁴ 1DFFB MODIFIER LETTER SMALL R WITH FISHHOOK AND MIDDLE TILDE. Figure 52.

Not requested:

The Sinological para-IPA letter η with tilde (Figure 22) is not requested at present due to questions over the appropriate Unicode name. I plan to submit a request when the naming issue of $\langle \eta \rangle$ and similar Karlgren letters is resolved.

Properties

The IPA beta, theta and chi are called LATIN rather than GREEK following U+019B LATIN SMALL LAMBDA WITH STROKE.

1DF3B;LATIN SMALL LETTER A WITH MIDDLE TILDE;Ll;0;L;;;;N;;;;
1DF3C;LATIN SMALL LETTER ALPHA WITH MIDDLE TILDE;Ll;0;L;;;;N;;;;
1DF3D;LATIN SMALL LETTER BETA WITH MIDDLE TILDE;Ll;0;L;;;;N;;;;
1DF3E;LATIN CAPITAL LETTER C WITH MIDDLE TILDE;Lu;0;L;;;;N;;;1DF3E;
1DF3F;LATIN SMALL LETTER C WITH MIDDLE TILDE;Ll;0;L;;;;N;;;1DF3D;1DF3D
1DF40;LATIN SMALL LETTER STRETCHED C WITH MIDDLE TILDE;Ll;0;L;;;;N;;;;
1DF41;LATIN SMALL LETTER INSULAR D WITH MIDDLE TILDE;Ll;0;L;;;;N;;;;

```
1DF42;LATIN SMALL LETTER ETH WITH MIDDLE TILDE;LI;0;L;;;;N;;;;
1DF43;LATIN SMALL LETTER SCHWA WITH MIDDLE TILDE;LI;0;L;;;;N;;;;
1DF44;LATIN SMALL LETTER G WITH MIDDLE TILDE;LI;0;L;;;;N;;;;
1DF45;LATIN SMALL LETTER SCRIPT G WITH MIDDLE TILDE;LI;0;L;;;;N;;;;
1DF46;LATIN SMALL LETTER H WITH MIDDLE TILDE;Ll;0;L;;;;N;;;;
1DF47:LATIN SMALL LETTER I WITH MIDDLE TILDE;LI:0;L:;;;N:;;;; + soft-dotted
1DF48;LATIN SMALL CAPITAL I WITH MIDDLE TILDE;LI;0;L;;;;N;;;;
1DF49;LATIN SMALL LETTER J WITH MIDDLE TILDE;Ll;0;L;;;;N;;;; + soft-dotted
1DF4A;LATIN SMALL LETTER K WITH MIDDLE TILDE;Ll;0;L;;;;N;;;;
1DF4B;LATIN SMALL LETTER L WITH PALATAL HOOK AND MIDDLE TILDE;Ll;0;L;;;;N;;;;;
1DF4C;LATIN SMALL LETTER L WITH RETROFLEX HOOK AND MIDDLE TILDE;LI;0;L;;;;N;;;;
1DF4D;LATIN SMALL LETTER L WITH BELT AND MIDDLE TILDE;LI;0;L;;;;N;;;;
1DF4E;LATIN SMALL LETTER LEZH WITH MIDDLE TILDE;Ll;0;L;;;;N;;;;
1DF4F;LATIN SMALL LETTER TURNED Y WITH MIDDLE TILDE;Ll;0;L;;;;N;;;;
1DF50;LATIN SMALL LETTER O WITH MIDDLE TILDE;Ll;0;L;;;;N;;;;
1DF51;LATIN SMALL LETTER PHI WITH MIDDLE TILDE;LI;0;L;;;;N;;;;
1DF52;LATIN SMALL CAPITAL R WITH MIDDLE TILDE;LI;0;L;;;;N;;;;
1DF53;LATIN SMALL LETTER R WITH TAIL AND MIDDLE TILDE;LI;0;L;;;;N;;;;;
1DF54:LATIN SMALL LETTER S WITH HOOK AND MIDDLE TILDE:LI:0:L:::::N:::::
1DF55:LATIN SMALL LETTER ESH WITH MIDDLE TILDE:LI:0:L::::N:::::
1DF56;LATIN SMALL LETTER TURNED T WITH MIDDLE TILDE;Ll;0;L;;;;N;;;;
1DF57;LATIN SMALL LETTER TESH DIGRAPH WITH MIDDLE TILDE;LI;0;L;;;;N;;;;
1DF58;LATIN SMALL LETTER THETA WITH MIDDLE TILDE;Ll;0;L;;;;N;;;;
1DF59;LATIN SMALL LETTER TURNED M WITH LONG LEG AND MIDDLE TILDE;LI;0;L;;;;N;;;;
1DF5A;LATIN SMALL LETTER U WITH MIDDLE TILDE;Ll;0;L;;;;N;;;;
1DF5B;LATIN SMALL LETTER UPSILON WITH MIDDLE TILDE;Ll;0;L;;;;N;;;;
1DF5C;LATIN SMALL LETTER V WITH MIDDLE TILDE;Ll;0;L;;;;N;;;;
1DF5D;LATIN SMALL LETTER V WITH HOOK AND MIDDLE TILDE;LI;0;L;;;;N;;;;
1DF5E;LATIN SMALL LETTER X WITH MIDDLE TILDE;LI;0;L;;;;N;;;;
1DF5F;LATIN SMALL LETTER CHI WITH MIDDLE TILDE;Ll;0;L;;;;N;;;;
1DF60;LATIN SMALL LETTER EZH WITH MIDDLE TILDE;Ll;0;L;;;;N;;;;
1DF61;LATIN SMALL LETTER GLOTTAL STOP WITH MIDDLE TILDE;LI;0;L;;;;N;;;;
1DF62:LATIN SMALL LETTER INVERTED GLOTTAL STOP WITH MIDDLE TILDE;LI:0;L;;;;N;;;;;
1DF63;LATIN SMALL LETTER N WITH DOUBLE MIDDLE TILDE;Ll;0;L;;;;N;;;;
1DF64;LATIN SMALL LETTER R WITH DOUBLE MIDDLE TILDE;Ll;0;L;;;;N;;;;
1DFFA;MODIFIER LETTER SMALL TURNED R WITH MIDDLE TILDE;Lm;0;L;<super> AB68;;;;N;;;;
1DFFB;MODIFIER LETTER SMALL R WITH FISHHOOK AND MIDDLE TILDE;Lm;0;L;<super> 1D73;
    ;;;N;;;;
```

Chart

Greyed out cells are assigned (medium grey) or proposed elsewhere (light grey).

1DF00			La	tin Ex	tend	1DFFF			
	1DF0	1DF1	1DF2	1DF3	1DF4	1DF5	1DF6		1DFF
0	fŋ	К	c	¥.	€	0	3		
1	9	ŀ	u	ħ	ъ	ф	2		
2	9	ಡ್ತ	σ	ф	ð	R	3	• • •	
3	k	ţ	ďз	q,	Э	ŧ	Ħ	• • •	
4	€	ŋ,	ţſ	Ŗ	g	ફ	¥		
5	ß	ત્ર	rd	ţ	g	∮			
6	Я	ß	ŀ	tş	ħ	3			
7	ű	뮟	'n	д	i	₩			
8	1	3₀	r	y	ł	₽			
9	f	dз	'n	X	j	щ			
Α	Į	į	t	Ş	ħ	¥			æ
В	f	ρ	β	a	ţ	ਲ			£
С	£	Ą	đ	a	t	¥			h
D	J	ત	dz	β	‡	Ж			ф.
E	Z	S	ð	€	抟	Ð			s,
F	C	T	9,	€	K	X			z

References

Hussain al Duwair *et al.* (2014) Las consonantes en español y árabe: un análisis contrastivo para fines didácticos. marcoELE. *Revista de Didáctica Español Lengua Extranjera*.

Muhammed Bakalla (1979) Arab and Muslim phoneticians. In Hollien (ed.) *Current Issues in the Phonetic Sciences: Proceedings of the IPS-77 Congress, Miami Beach, Florida, 17–19th December 1977*, part 1.

Matthias Castrén (1854) Grammatik der Samojedischen Sprachen. St. Petersburg.

(1857) Versuch einer burjätischen Sprachlehre. St. Petersburg.

(1858) Versuch einer ostjakischen Sprachlehre. St. Petersburg.

J. C. Catford (1977) Fundamental problems in phonetics. Indiana University Press.

Peter Daniels (2014) The type and spread of Arabic script. In Mumin & Versteegh (eds.) *The Arabic Script in Africa*, pp 25–39.

Olivier Durand (2017) Remarques sur le consonantisme de l'arabe. Folia Orientalia, vol. 54.

Maoji Fu (1997) A descriptive grammar of Lolo. Linguistics of the Tibeto-Burman Area, 20.1.

Edward Lipiński (2001) Semitic Languages: Outline of a Comparative Grammar.

John Kelly & John Local (1989) Doing Phonology. Manchester University Press.

Alexei Kassian, ed. (2020) 110-item wordlists for the Khinalug group (North Caucasian family). Moscow lexical database. http://moslex.org/ (accessed: 28 Sep 2023)

Geoffrey Khan (2015) Domains of Emphasis, Syllable Structure and Morphological Boundaries in the Christian Urmi Dialect of Neo-Aramaic. In Khan and Napiorkowska, eds., *Neo-Aramaic and its Linguistic Context*. Piscataway: Gorgias. p. 145–161.

Arne Krause (2018) Transkription Deutsch-Arabischer Diskurse. MuM-Multi II, TP Hamburg.

Roger Lass (1984) Phonology: An Introduction to Basic Concepts. Cambridge University Press.

Unn Gyda Næss (2008) "Gulf Pidgin Arabic": Individual strategies or a structured variety? A study of some features of the linguistic behaviour of Asian migrants in the Gulf countries. University of Oslo.

Master's thesis.

Cathair Ó Dochartaigh (1997) Survey of the Gaelic dialects of Scotland. Dublin Institute for Advanced Studies.

Kenneth Pike (1947) Phonemics. University of Michigan Press.

Sibylle Reichel (April 2003) Handbuch zum Zeichensatz SMFTeuthonista.

Cynthia Shuken (1980) *Instrumental investigation of some Scottish Gaelic consonants*. University of Edinburgh.

A. N. Tucker (1946) Foreign sounds in Swahili. *Bulletin of the School of Oriental and African Studies* 11:4, p. 854–871.

The UCLA Phonetics Lab Archive (2007) Arabic, Tunisian Spoken. UCLA Department of Linguistics.

UPSID (1981): UCLA Phonological Segment Inventory Database.

Upton, Parry & Widdowson (1994) Survey of English Dialects. Routledge.

Manfred Woidich (2006a) Das Kairenisch-Arabische: eine Grammatik. Harrassowitz Verlag.

Manfred Woidich (2006b) Cairo Arabic. In Versteegh et al., *Encyclopedia of Arabic Language and Linguistics*, vol. 1. Brill.

Figures

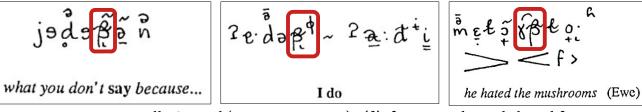
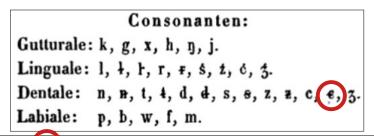


Figure 1. Kelly & Local (1989: 135–136, 69). $\langle \beta \rangle$ for Tyneside English and for Ewe.

We refer to these resonance categories as palatalised, clear, half-clear, central, half-dark, dark and velarised. Using C as a symbol for any consonantal symbol we notate them using diacritics thus:

Figure 2. Kelley & Local (1989: 73). Capital $\langle \mathfrak{C} \rangle$ as a wildcard for a velarized consonant. $\langle \mathfrak{C} \rangle$ for a palatalized consonant is supported at U+A7C4. (K&L make a semantic distinction in the horizontal placement of the tilde, but this is not IPA usage and is not requested here.)



§ 39. s ist ein mouillirtes c, welches man im Russischen durch ць ausdrücken könnte, im Deutschen durch zj, im Finnischen u. s. w. durch tsj. Dieser Laut kommt in keiner andern Samojedi-

Figure 3. Castrén (1854: 2). $\langle \epsilon \rangle$ for Samoyed. This represents an affricate, [tsⁱ]. That is, the tilde indicates palatalization. Note that the rather obscure letter $\langle f \rangle$ is supported at U+1DF11.

2 a; h - Σ H - γ X - k G - D j / y/i:
- 1 r n - T d t - S z s - (δ) δ θ f b m w/u:

The symbols used above represent certain Arabic sound units.

1 stands for the voiced pharyngeal fricative; H for the voiceless pharyngeal fricative; X for the voiceless uvular fricative; the voiced uvular fricative; G the voiced uvular stop; the capital letters represent the emphatic counterparts of the non-emphatics; the glottal stop; the emphatic counterpart of the voiceless interdental fricative; the colon after the short vowels indicates vowel lengthening. I is a voiceless palatoalveolar fricative. I is a voiced fricative lateral emphatic (see the following diagram which was made by Al-Sakkaki in the 13th century).

Figure 4. Muhammed Bakalla (1979: 10). $\langle \delta \rangle$ is used for Arabic $\not = /\delta^c/$, the emphatic partner to $\not = /\delta/$. The symbol for the emphatic lateral fricative (yellow) is difficult to discern and is not requested; it might be a script el $\langle \pounds \rangle$. Standard IPA would be $\langle \frac{1}{5} \rangle$.

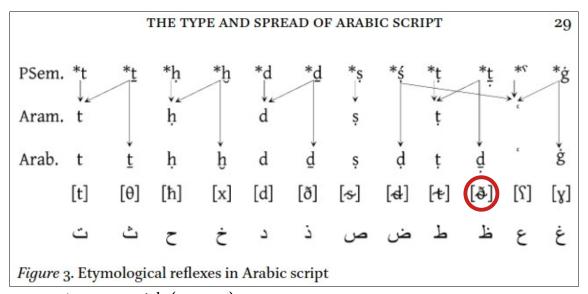


Figure 5. Daniels (2014: 29).

La mayor diferencia entre las consonantes fricativas árabes y las españolas está en la elevada presencia de fricativas del árabe, que se articulan en posición post-velar (dos uvulares, dos faríngeas, una glotal), así como en la presencia de /ð/y /s/, fricativas enfáticas. El AEM cuenta con catorce consonantes fricativas de las cuales dos son enfáticas; la /ð/ y la /s/, opuestas a sus correspondientes no enfáticas; la /ð/ y la /s/. Sustituir uno de estos fonemas enfáticos por otro no enfático es suficiente para que se altere y pase a evocar otro significado. Este fenómeno se comprueba en pares mínimos como

Figure 6. Al Duwair (2014: 2).

		Bilabial	Labiodental	Interdental	Pharyngealized ¹	Dental	Alveolar Pharyngealized	Alveolar	Alveo-Palatal	Velar	Uvular	Pharyngeal	Glottal
Stops	VL VD	b [b]				t [t̪] d [d̪]	ţ[ŧ]			k [k] g [g]	q [q]		,[3]
Fricatives	VL VD		f [f]	t [θ] d [ð]		s [s] z [z]	ș [s]		š [ʃ]	ḫ[x] ġ[ɣ]		, [¿] , [¥]	h[h]
Affricates	VL VD				z [ð]			č [tʃ] j[d3]					
Nasals	VD	m [m]						n[n]					
Тар	VD							r[r]					
Approxim ants	VD							1[1]	y [j]	w[w]			

Figure 7. Næss (2008).

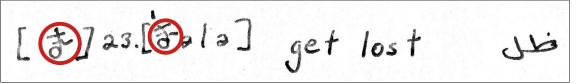


Figure 8. The UCLA Phonetics Lab Archive (2007).

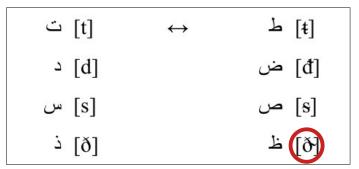


Figure 9. Durand (2017: 125).

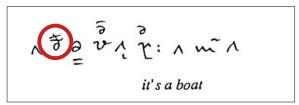


Figure 10. Kelly & Local (1989: 142). (8) for Malayalam.

Dans un cas comme dans l'autre, il y a perte du phonème / ف de l'arabe ancien. Exceptions connues jusqu'à ce jour:

- Jabal Yazīdī, Sud du Yemen, qui distingue calirement d [đ] de o [ð], en accord avec la norme canonique;
- quelques points du Nord du Yemen, qui distinguent les deux phonèmes comme $\check{\rho}([\check{\delta}])$ pour $\dot{\psi}$ e $\check{\rho}([\theta])$ sourd!) pour $\dot{\psi}$;

Figure 11. Durand (2017: 145). $\langle \eth \rangle$ and $\langle \Theta \rangle$ for Arabic.

Figure 12. Gairdner (1925: 12). An insular $\langle b \rangle$ and b-tilde $\langle b \rangle$. These forms are also found in Tucker (1946). The twisted tail of $\langle g \rangle$ is typical of the period, and was the form found on the 1912 IPA chart.

2) Lockerung des Verschlusses

Ein Übergang zwischen Verschluß- und Reibelaut wird mit einer Tilde durch die Ober- bzw. Unterlänge des Verschlusslaut-Zeichens markiert.

b zwischen b und w

g

zwischen d und einem entsprechenden Reibelaut

zwischen g und x

Figure 13. Reichel (2003: 8). Single-bowl G with tilde, $\langle q \rangle$.

Exercise 2. Repeat for pharyngeal

Figure 14. Pike (1947: 33). Double-bowl G with tilde, $\langle g \rangle$, for a voiced pharyngeal stop.

g g before front
vowels)

k (k before front
vowels)

x (xy before back
vowels)

k: (k before front
vowels)

kinch before front
vowels)

k' (k' before front
vowels)

Figure 15. Kassian (2020). $\langle g \rangle$, $\langle t \rangle$ and $\langle x \rangle$ in transcription system for Khinalug.

§ 1. Zur Bezeichnung der Laute der ostjakischen Sprache werden in vorliegender Arbeit folgende Buchstaben angewandt:

Vocale: a, e, i, o, ö, u, ü.

Consonanten: k k g, g, x, ŋ, j, l, l, r, ś, ć, ź, n, n, t, t, t, t, d, d, d, d, d, e, s, c, z, p, b, w, m.

5) g, k mit d, t, z. B. nîngem, nîndem, ruhen; kenak tenak, Warze.

Figure 16. Castrén (1858: 1, 20). $\langle g \rangle$ and $\langle h \rangle$ for Khanty.

bei emphatischer Umgebung: sulb ['sulb] "Stahl"; šunat ['funat] [u] "Taschen"; duhr [duhr] "Nachmittag".

Figure 17. Woidich (2006a: 8). Phonetic $\langle f \rangle$, $\langle h \rangle$ and $\langle u \rangle$ (in the last example duhr ['duhr]) in the environment of emphatic consonants in Arabic.

Word-initial strings with decreasing sonority such as npalla are often pronounced with epenthetic prosthetic vowels, which is evidence for this analysis of the syllable structure, e.g.: nšákla [ɪnˈʃɪk̪lɪ] 'he kissed him' +nţúrrə [unˈturrə] 'guard it' [it|didli] 'it trembled' rjádla +rxátla [ərˈxətlə] 'he ran'

+háva [ˈhaːva·] 'air, weather' < Pers. (< Ar.) hăvấ CVV.CVV

Figure 18. Khan (2015: 148, 154). $\langle u \rangle$, $\langle \vartheta \rangle$ and $\langle x \rangle$ for Neo-Aramaic: ntúrra[un'turra], $rx\acute{a}tla$ [ər'xətlə]; $\langle h \rangle$, $\langle a \rangle$ and $\langle v \rangle$ in $h\acute{a}va$ ['ha:va']. Overstruck tilde U+0334 is recoverable from the PDF where it is not particularly visible here.

Das /l/ des Artikels assimiliert sich auch an emphatische Konsonanten, das /i/ kann daher emphatisch werden: ittama [i t:amas] issubh [i's:ubħ] "die Gier" "der Morgen" irrāgil [i ˈfːaːgil] "der Mann" iggār [i g:a:r] "der Nachbar" aber: innahār [in:a ha: F] "der Tag" aber mit Ausbreitung: yidrab ['jidrab] titxattat [tit 'xat:at] "er schlägt" "sie plant" yirxaslak [jif 'xaslak] til'it ['til'it] "es ist billiger für dich" "sie stieg" sannitha [saˈnɨːtha] yixbat ['jexbat] "ihr Gestank" "er schlägt" [kë 'tëpth] b + t[pt] katabt "ich schrieb"

b + s"er druckt ab" > [ps] yibsum ["jæpsum] "Körper pl." /g + s/agsām > [ks] [?ak'së:m]

Figure 19. Woidich (2006a: 26, 18). The "emphatic" vowels $\langle i \rangle$, $\langle i \rangle$, $\langle i \rangle$ and possibly $\langle u \rangle$, though the last is not clear. (Also $\langle q \rangle$ and $\langle k \rangle$.) "Emphatic" a (as in titxattat) however is transcribed $\langle a \rangle$.

When inflectional or derivative affixes are attached to the end of phonologically integrated loanwords the consonants retain their voicing, e.g.:

```
'avádva [?a'vidva]'he used to do'barázva [ba'rizva]'it used to dry'*xiyavánda [xi'ja:vandə]'roads'*kassába [kas'sa:bə]'butchers'*'otága [?a:'ta:və]'rooms'*sarbázan [sar'ba:zan]'our solider'
```

Figure 20. Khan (2015: 157). Tildes for emphatic sounds in Neo-Aramaic, including vowels: $\langle \mathbf{e} \rangle$, $\langle \mathbf{x} \rangle$ and $\langle \mathbf{v} \rangle$ in xiyavándə [xiˈjɑːvɑndə]; schwa $\langle \mathbf{e} \rangle$ in kassábə [kɑsˈsɑːbə], $\langle \mathbf{e} \rangle$ and $\langle \mathbf{f} \rangle$ in 'otáġə [ʔoːˈtɑːyə], and $\langle \mathbf{e} \rangle$ in sarbázan [sɑrˈbɑːzɑn]. The last is typeset with a barely visible U+0334 tilde overlay on the a ($\langle \mathbf{a} \rangle$) in the PDF.

```
This applies also to loanwords, e.g.:

+ xiyávand [xi'ja:vanth] 'road' (Pers.)

+ kássab [ˈkassaph] 'butcher' (Pers. < Ar.)

+ časəb [ˈtʃasəph] 'poor' (Pers. < Ar.)
```

Figure 21. Khan (2015: 157). $\langle \mathfrak{g} \rangle$, $\langle \mathfrak{a} \rangle$ and $\langle \mathfrak{d} \rangle$ for Neo-Aramaic, in časəb [\mathfrak{g} asəph].

	Front
Distance	
Very close	1 ï (3) 1 ï

(iii) — is a velarized γ . It can be acquired by pronouncing the vowel γ and at the same time raising the root of the tongue toward the soft palate. The lips are widely *spread*. For example: $\sqrt[3]{s}\gamma$ to know.

```
<sup>2</sup>nw<sup>3</sup>sη <sup>2</sup>sη? 'Do you know?'

<sup>2</sup>nw <sup>3</sup>γ <sup>2</sup>γ? 'Do you sell?'

<sup>2</sup>nw <sup>2</sup>ts'η <sup>3</sup>he <sup>3</sup>γ <sup>2</sup>γ? 'Do you like it?'
```

Figure 22. Fu (1997: 56, 60, 62). $\langle \gamma \rangle$ for Lolo. Although typically labeled a "vowel," $[\gamma]$ lies beyond normal vowel space and is arguably a syllabic consonant, similar to the syllabic $[\gamma]$ also seen here. A Unicode character is not requested for now due to the ongoing discussion over disunifying the Karlgren letter $\langle \gamma \rangle$.

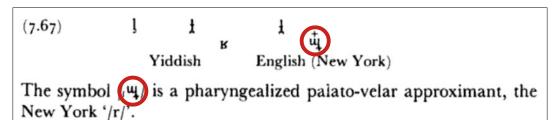


Figure 23. Lass (1984: 158). (u) for New York English.

Spread of emphasis: suprasegmental spread of pharyngealization starting from an emphatic consonant takes place regularly as in *baṭnaha* [baṭnaha] 'her stomach', *taṣliḥāt* [ṭaṣlɨħaɪṭ] 'repairs', *abyad* [ʔabɨad] 'white'.

Figure 24. Woidich (2006b: 325). $\langle j \rangle$ and $\langle i \rangle$ for Arabic.

```
p = [\varphi] (do. spirantized)

q = [k], [k'], [k] (emphatic voiceless velar plosive, glottalized or velarized)

q^y = [k] (palatalized and glottalized voiceless velar plosive)
```

Figure 25. Lipiński (2001: 103). $\langle k \rangle$ for Arabic q. Though the glyph is not clear in this image, all the 'velarized' consonants in the chart are marked with a tilde.

2.13. Si */ĝ/ devient [k²] dans les lang la pharyngalisation en [k] se révèle difficile dans l'articulation de la vélaire simple [k], le

Figure 26. Durand (2017: 143).

- [q] postvelar, Verschlußlaut, stimmlos; marginal, in Bildungswörtern aus dem Hocharabischen: ilQur'ān "der Koran"; inqilāb "Umsturz"; qarya "Dorf"; raqam "Nummer". Nicht selten wird /q/ als velarisiertes [k] = [k] gesprochen: iktiṣadiyya "wirtschaftlich"; gabal ilmukaṭṭam (= ilmuqaṭṭam) "der Muqaṭṭam-Berg". Zu humoristischen Zwecken kann es /k/ ersetzen: sibti f-Maṣri ḥittit dīn qaṭqūṭa (katkūta) "ich habe in Kairo 'ne dufte Biene zurückgelassen" [ME].
- /t/ [t] dental, Verschlußlaut, stimmlos, emphatisch: tifl "Kind"; baṭal "Held"; balāt "Fließen".
- /k/ (k) velar, Verschlußlaut, stimmlos, emphatisch; als Ersatz für postvelares /q/ in Bildungswörtern: iktiṣād "Ökonomie".

Figure 27. Woidich (2006a: 12).

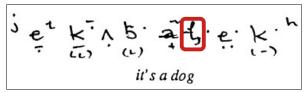


Figure 28. Kelly & Local (1989: 122). ⟨⅓⟩ for Sinhalese.

barley-ails n barley AWNS II.5.3. ba'[ti-e'utz O, batte-eiutz Ha/while growing], batte-eitz Ha, batte-eitz W, batte-eitz W, batte-eitz He, batte-aitz Ha, batte-aitz D Ha, batte-aitz Ess, batte-aitz So W, batte-aitz W, batte-aitz So, batte-aitz Wo[old], batte-aitz Sr, batte-bitz D, batte-bitz Ha/after threshing], batte-bitz Co, batte-bitz O

Figure 29. Upton, Parry & Widdowson (1994: 26). $\langle \uparrow \rangle$ contrasting with both $\langle \uparrow \rangle$ and $\langle \mid \rangle$ in English dialects.

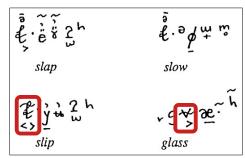


Figure 30. Kelly & Local (1989: 38). $\langle \hat{\uparrow} \rangle$ and $\langle \psi \rangle$ for disordered speech in English.

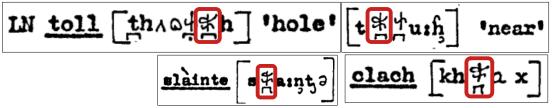


Figure 31. Shuken (1980: 126, 151, 223, 233). Dental [*] in Scots Gaelic.

(iṭṭalaȝa «monter» < *iṭ-ta-laȝa): pour *iḍ-ta-raba on s'attendrait à *iṭṭaraba ou *iddaraba, et non à إِضْطَرَبُ iḍṭaraba; force est de postuler un ḍ fricatif, '[ilʒtaraba].

them. The air goes out through the side(s) passage of the mouth». Soit (\dot{z}) à transcrire plus précisément (\dot{z}) . Concrètement: mettre la langue en position pour [1], faire [z], en pharyngalisant le tout.

Figure 32. Durand (2017: 146). 〈读〉 for Arabic. Cf. the *ad hoc* transcription of this sound in Figure 4.

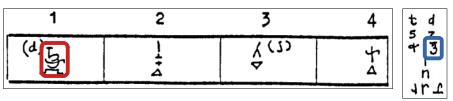


Figure 33. Shuken (1980: 252). Interdental [\S] in Scots Gaelic. The triangle and inverted triangle for apical and laminal articulation; the current IPA diacritics $\langle \ \rangle$ and $\langle \ \rangle$ only date to the Kiel Convention of 1989. Shuken's somewhat idiosyncratic handwritten variant of $\langle \ \rangle$ (blue) is identified as such in the alveolar column of the consonant table reproduced at right (from p. 34).

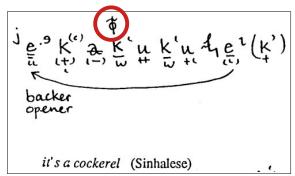


Figure 34. Kelly & Local (1989: 85). $\langle \varphi \rangle$ for Sinhalese. The vertical stacking of full-size $\langle \varphi \rangle$ and $\langle \underline{k}' \rangle$ indicates that these are two phonetic realizations $\{\varphi, \underline{k}'\}$ recorded for that consonant.

r as in ramh or barr, in e.g. the Outer Hebrides, is written with capital to indicate the rolling, and the velarization mark. The corresponding velarized fricative is [x]. The lenited 'broad' r is

Figure 35. Ó Dochartaigh (1997: 130). $\langle \mathbb{R} \rangle$ for the Otter Hebrides dialect of Scottish Gaelic. This is not the IPA uvular $\langle \mathbb{R} \rangle$, as the small capital simply indicates a trill.

	1	2	3
L	47.5	1	₹
N	47\$	774	يرار دياره
R	T	<u>r~r</u>	あ る

Figure 36. Shuken (1980: 252). $[\mathfrak{x}]$ in the Bernera and Ness dialect of Scottish Gaelic. The tilde is reversed, but this is simply a manuscript variant (the nasal tilde is the same: see $[\tilde{\mathfrak{x}}]$ in next figure and in the key in Shuken 1980: 34).

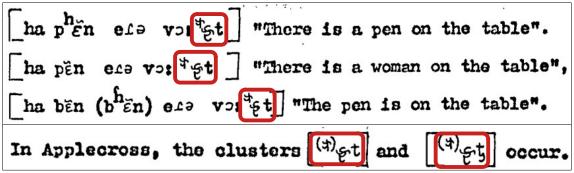


Figure 37. Shuken (1980: 74, 81). $\langle \S \rangle$ and modifier $\langle * \rangle$ in the Applecross dialect of Scottish Gaelic. [* \S] is the realization of /r/ in this environment, [vo:* \S t] being orthographic *bhòrd*.

Velarized voiceless palato-alveolar sibilant fricative. /f/ 1 Russian.

Figure 38. UPSID (1981: 203). $\langle f \rangle$ for Russian.

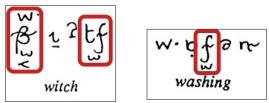


Figure 39. Kelly & Local (1989: 154, 252). $\langle f \rangle$ (and $\langle \beta \rangle$) for a child's English.

$$/t + \dot{g}/ > [dy]$$
 yitġadda [jɪdˈɣɛ̈dːɛ̃] "er ißt zu Mittag"
/f + d/ > [vel] yifḍal [jɨvelαł] "er bleibt"

Figure 40. Woidich (2006a: 17). Phonetic ⟨⟨⟨⟩ in Arabic.

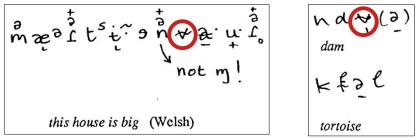
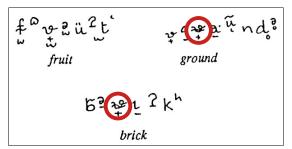


Figure 41. Kelly & Local (1989: 90, 150). Onset $[\psi]$ and syllabic $[\psi]$ in Welsh and Fang.



instance. In impressionistic records made at the beginning of the period of study these words appear with word initial configurations such as

That is, we find an exo-labial glide occurring with or without dark resonance. Some three and a half months later the same words are only

Figure 42. Kelly & Local (1989: 256). ⟨₺⟩ for child-language acquisition of English.

§ 2. Die sowohl allen Dialekten gemeinsamen als auch einzelnen derselben eigenthümlichen Laute, welche letztern mit einem Sternlein (*) bezeichnet sind, lassen sich also zusammenfassen:

Vocale: a, *ä, e, o, ö, u, u.

Consonanten: k, *k, x, *x, h, g, ŋ, j, l, ł, r, ғ, ś, ź, *ć, *ȝ, n,
*n, t, *ŧ, d, *d, s, z, *c, ȝ, p, b, m.

§ 11. Was die Consonanten anbetrifft, so ist zu bemerken:

1) * (x,)1, +, +, +, + sind Mouillirungen von k, x, l, r, n, t, d.

moken NU., moken(n) T., moke S., moken moken mokin Ch., Kauharz;

Figure 43. Castrén (1857: 1, 3, 174). $\langle x \rangle$ (and $\langle t \rangle$) for Buryat.

are sometimes pharyngealized. Pharyngealized (x) and (xw) occur in the Bzyb dialect of Abkhaz, in contrast with both plain uvulars and

soft palate. In χ there is a fricative-like channel extending from the rear velar zone, over the uvula and down into the radico-pharyngeal zone. In Bzyb, incidentally, 'extended' uvular-pharyngeal χ contrasts with both $[\chi]$ and [h], as in $[a'\chi \ni]$ 'head', $[a'\chi \ni]$ 'lead' (n.), $[a'haw \ni]$ 'air'.

Figure 44. Catford (1977: 193, 195). ⟨χ⟩ for Abkhaz.

Velarized voiced palatal lateral approximant. //// 1 Irish.

Figure 45. UPSID (1981: 219). $\langle f \rangle$ for Irish. Some care needs to be taken graphically to distinguish this from a voiceless fricative $\langle f \rangle$. The sound is also reported for Arabic, e.g. in Durand (2017), but so far we've only found it transcribed as $\langle f \rangle$, with the subdot convention for emphatic consonants.

Velarized voiced palato-alveolar sibilant fricative. Russian.

Figure 46. UPSID (1981: 203). ⟨₹⟩ for Russian.

Beschreibung	Lautschrift IPA
nicht gepresster Stimmlippenverschlusslaut (Kehlkopfverschlusslaut)	[2]
kommt auch velarisiert vor, wie z.B. im Wort الله (DMG: Aḷḷāh [ʔaʾłːaːh], Allah, Gott'	(?)

Figure 47. Krause (2018: 3). $\langle 2 \rangle$ and $\langle \hbar \rangle$ for Arabic.

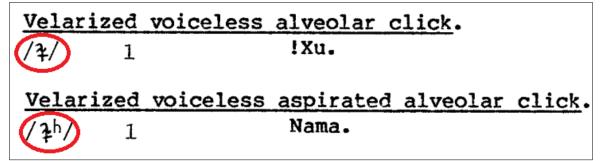


Figure 48. UPSID (1981) p. 187. [Intervening lines omitted.] The velarized click $\langle \mathfrak{z} \rangle$ in !Xu and Nama. Because the IPA at the time had only two letters – $\langle \mathfrak{z} \rangle$ and $\langle \mathfrak{c} \rangle$ – for the three articulations of dental, alveolar and palatal clicks, Maddieson used $\langle \mathfrak{z} \rangle$ for dental, $\langle \mathfrak{z} \rangle$ for alveolar, and $\langle \mathfrak{c} \rangle$ for palatal.

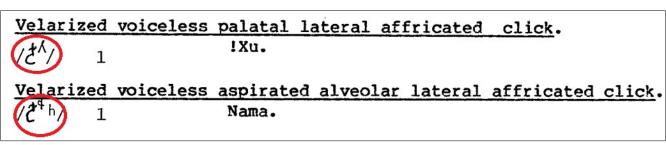


Figure 49. UPSID (1981: 190). [Intervening lines omitted.] The velarized click $\langle \mathring{3} \rangle$ in !Xu and Nama. If the author were accostomed to transcribing clicks with the pipe letters $\langle | \parallel ! \nmid \rangle$, and only switched to the official IPA letters for this document, that might explain why this letter got reversed.

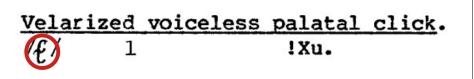


Figure 50. UPSID (1981: 188). The velarized click $\langle \xi \rangle$ in !Xu. (There is no example from Nama here, because the corresponding consonant is transcribed as palatoalveolar via a retraction diacritic, [\mathfrak{z}^h].)

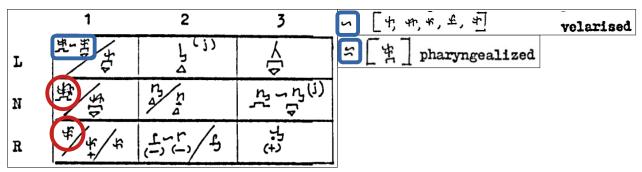


Figure 51. Shuken (1980: 253, 34–35). $\langle \frac{1}{7} \rangle$, $\langle \frac{1}{8} \rangle$ and $\langle \frac{2}{7} \rangle$ in the Outer Hebrides dialect of Scottish Gaelic. A single tilde indicates velarization and a double tilde indicates pharyngealization (see key to diacritics at right), following the IPA convention of doubling a diacritic to indicate a greater degree of that feature (pharyngealization being more "guttural" than velarization). Pharyngealized [1], but not [1] or [1], is also recorded from Skye and Ross-shire.

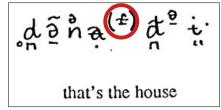


Figure 52. Kelley & Local (1989: 163). A superscript letter ⟨+⟩ for Welsh.

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 dunimintunativo

A. Aummstrative						
1. Title:	Additio	nal phonetic click letters				
2. Requester's nam	ne:	Kirk Miller				
3. Requester type (Member body/Liaison/Individual contribut	ion): individual				
4. Submission date		2023 October 13	3			
5. Requester's refe	rence (if applicable):					
6. Choose one of th						
	omplete proposal:		yes			
(or) More	information will be provided later:					
B. Technical – Ger	neral					
1. Choose one of th						
	osal is for a new script (set of characters):					
	osed name of script:					
	osal is for addition of character(s) to an exist	ing block:	<u>yes</u>			
Name	e of the existing block:	Latin Extended-G				
2. Number of chara	acters in proposal:		44			
3. Proposed catego	ory (select one from below - see section 2.2 o	f P&P document):				
A-Contemporar			ction)			
C-Major extinct		E-Minor extinct				
	oglyphic or Ideographic	G-Obscure or questionable usage sy	mbols			
4. Is a repertoire in	ncluding character names provided?		yes			
	the names in accordance with the "charact	er naming guidelines"				
	nex L of P&P document?		yes			
b. Are the ch	naracter shapes attached in a legible form su	itable for review?	yes			
5. Fonts related:	1					
	provide the appropriate computerized font t	o the Project Editor of 10646 for publishing	ng the standard?			
W *** P		k Miller	ag viie sourium ur			
b. Identify th	ne party granting a license for use of the fon		ftp-site, etc.):			
		tium Release)	1 , ,			
6. References:						
	ences (to other character sets, dictionaries, d	lescriptive texts etc.) provided?	ves			
	shed examples of use (such as samples from					
sources)	new enumpres of use (euch us sumpres it em	ne mep up er e, muguzmee, er etmer				
,	characters attached?	yes				
7. Special encoding						
	oposal address other aspects of character dat	ta processing (if applicable) such as input.				
	, sorting, searching, indexing, transliteration					
presentation	., 001 01119, 0001 011119, 11100119, 111011001	on ever (in) to produce energed initial industrial				
8. Additional Infor	mation:					
	ited to provide any additional information a	shout Properties of the proposed Characte	ar(s) or Script that			
	ct understanding of and correct linguistic p					
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						
	/reports/tr44/) and associated Unicode Tec					
	Committee for inclusion in the Unicode Sta		,			

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.)?	<u>yes</u>
If YES, with whom? The authors are members of the user community.	
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)	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 e	ntirely
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)?	if possible
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?	<u>yes</u>
If YES, is a rationale for its inclusion provided?	<u>yes</u>
If YES, reference: Unicode proposal N2632 and its acceptance in Unicode 4	.0.
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?	<u>no</u>
If YES, are the equivalent corresponding unified ideographic characters identified?	
If YES, reference:	