L2/21-207

Unicode request for Latin subscript letters

Kirk Miller, kirkmiller@gmail.com 2021 August 16

Semantically distinct subscripts are used in several phonetic traditions. These include informal extensions of the IPA, the RFE (*Revista de Filología Española*), and Americanist phonetic notation, as well as morphophonemic transcription in the eastern Slavic tradition.

Among the lower-case letters of the basic Latin alphabet, Unicode currently lacks subscript *b c d f g q w y z*. All but *q* are attested in this request. Of Greek, Unicode currently supports subscript *β γ ρ φ χ*, and the only additional IPA letter is the schwa (ₔ). Several unsupported subscript IPA letters are illustrated in the figures below, but the majority of the IPA is unattested.

In an extension of official IPA transcription, some phoneticians use subscript modifier letters for simultaneous articulation and restrict the official superscript modifiers to sequential sounds, such as onsets, on/off-glides, releases and weakly or incompletely articulated consonants. For example, Penhallurick (1991) uses subscript *ɹ ʕ ʁ ʁ* for r-coloring of a preceding vowel, contrasting with the corresponding superscripts, which indicate a weakened rhotic consonant.

For Americanist notation, the American Anthropological Association (1916: 9) says, "If it is desired to distinguish between vocalic timbres and weakly articulated voiceless vowels on the one hand and vocalic glides and weakly articulated voiced vowels on the other hand, superior vowels (ᵃ) may be used for the former, inferior vowels (ₐ) for the latter." The Americanist subscript forms that are not yet supported by Unicode are open *ɔ* and Greek *α ε ι υ ω* (ₔ ₐ ₑ ᵢ ᵤ ₒ).

In the east Slavic tradition, subscripts are used for phonemic conflation (archiphonemes). This convention is found in both Latin and Cyrillic script; Cyrillic characters are illustrated in L2/21-107 *Unicode request for Cyrillic modifier letters*. An illustration of both scripts is presented in Figure 1:

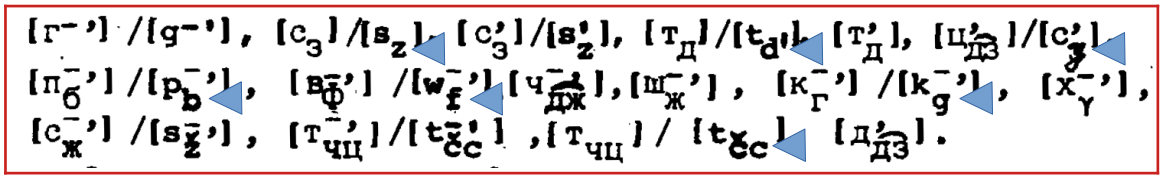


Figure 1. Kalnyn' & Maslennikova (1981: 396). Equivalent Russian (Cyrillic) and Polish (Latin) archiphonemes. Not all Russian archiphonemes have a Polish equivalent. The Latin subscripts include *b c d f g z ʒ*; subscript *č ž* should be handled with a combining diacritic. [The hand-written subscript *ʒ* [dz] is illegible here, but is identifiable from the figures below.] The diacritics indicate palatalization: non-palatalized ⟨C⁻⟩ vs palatalized ⟨C'⟩. ⟨C⁻'⟩ covers both.

In this tradition, subscripts indicating the approximate range of an archiphoneme contrast with superscripts indicating precise phonetic values, as in Figure 2:

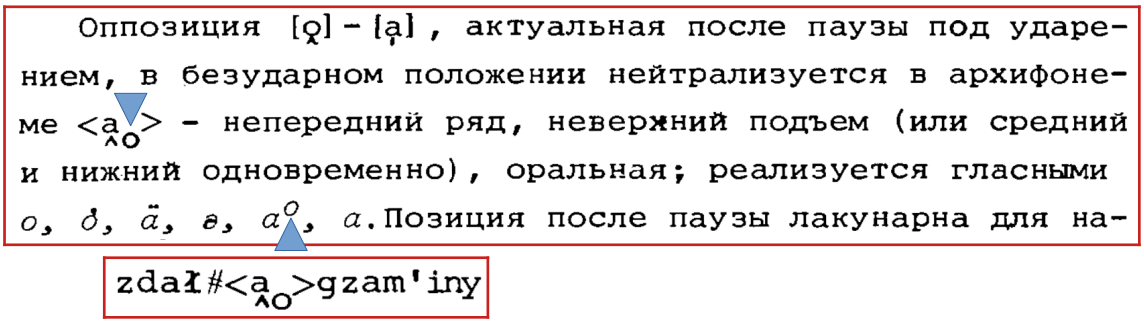


Figure 2. Kalnyn' & Maslennikova (1981: 163). An illustration of a semantic distinction between superscript and subscript modifier letters. Here the opposition of Polish [ǫ̨] and [ą̇], which is found in stressed post-pausa position, is neutralized to the archiphoneme ⟨a͡o⟩ in unstressed position. ⟨a͡o⟩ is realized phonetically by the vowels *o, ȯ, ä, ə, aᵒ, a*, thus contrasting the archiphonemic subscript notation a_o with the phonetic superscript notation a^o. That these are Latin letters can be seen by the several Polish examples, such as *zdał#⟨a͡o⟩gzam'iny* shown here.

In the early 20th century, publications were often typeset with some letters in small type, as seen in Figure 3 and in the illustration of the Spanish RFE in Figure 11. These small letters are typographic variants of subscripts and should generally be encoded as such.

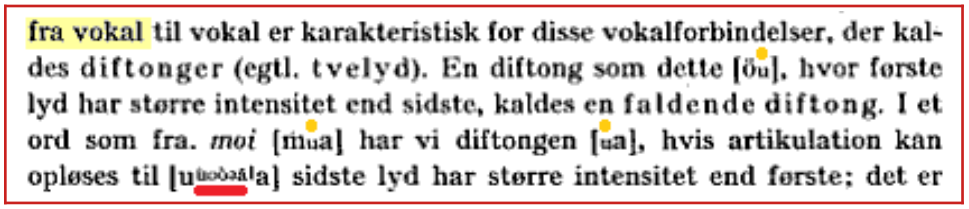


Figure 3. Hammerich (1934: 59), with weak elements of diphthongs printed in small script (yellow), as was common at the time. This contrasted with superscript (red).

In SIL fonts, most Americanist combining diacritics work properly on subscripts, e.g. superscript dieresis ⟨ä a̤⟩, dot ⟨ȧ a̤⟩, vertical stroke ⟨a̍ a̍⟩ and hacek ⟨š s̬⟩, and subscript circumflex ⟨a̰ a̰⟩, stroke ⟨a̵ a̵⟩ and ring ⟨ḁ ḁ⟩, though not the ogonek: ⟨ą ą̣⟩.

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

Chart

I propose adding Latin modifiers to the end of Latin Extended-G, to avoid mixing modifier and non-modifier letters in the block.

	…0	…1	…2	…3	…4	…5	…6	…7	…8	…9	…A	…B	…C	…D	…E	…F
Latin Extended-G																
U+1DFEx																ᵬ
U+1DFFx	ᵭ	ᵮ	ᵯ	ᵰ	ᵱ	ᵲ	ᵳ	ᵴ	ᵵ	ᵶ	ᵷ	ᵸ	ᵹ	ᵺ	ᵻ	ᵼ

Characters

ᵇ 1DFEF LATIN SUBSCRIPT SMALL LETTER B. Figures 1, 4.

ᶜ 1DFF0 LATIN SUBSCRIPT SMALL LETTER C. Figures 1, 5–8.

ᵈ 1DFF1 LATIN SUBSCRIPT SMALL LETTER D. Figures 1, 9–10.

ᵭ 1DFF2 LATIN SUBSCRIPT SMALL LETTER D WITH STROKE. Figure 11.

ᵋ 1DFF3 LATIN SUBSCRIPT SMALL LETTER OPEN E. Figure 12.

ᶠ 1DFF4 LATIN SUBSCRIPT SMALL LETTER F. Figures 1, 13.

ᵍ 1DFF5 LATIN SUBSCRIPT SMALL LETTER G. Figures 1, 14.

ᵞ 1DFF6 LATIN SUBSCRIPT SMALL LETTER GAMMA. Figures 15–17.

ᵩ 1DFF7 LATIN SUBSCRIPT SMALL LETTER PHI. Figure 33.

ᵣ 1DFF8 LATIN SUBSCRIPT SMALL LETTER TURNED R. Figures 18, 20.

ᵨ 1DFF9 LATIN SUBSCRIPT SMALL LETTER R WITH TAIL. Figures 18–20.

ᵣ 1DFFA LATIN SUBSCRIPT SMALL CAPITAL R. Figures 18–19.

ᵲ 1DFFB LATIN SUBSCRIPT SMALL CAPITAL INVERTED R. Figures 18, 21.

ʷ 1DFFC LATIN SUBSCRIPT SMALL LETTER W. Figures 22–23.

ʸ 1DFFD LATIN SUBSCRIPT SMALL LETTER Y. Figures 24–25.

ᶻ 1DFFE LATIN SUBSCRIPT SMALL LETTER Z. Figures 1, 26–30.

ᶾ 1DFFF LATIN SUBSCRIPT SMALL LETTER EZH. Figures 1, 31–32.

Properties

1DFEF;LATIN SUBSCRIPT SMALL LETTER B;Lm;0;L;<sub> 0062;;;;N;;;;;
1DFF0;LATIN SUBSCRIPT SMALL LETTER C;Lm;0;L;<sub> 0063;;;;N;;;;;
1DFF1;LATIN SUBSCRIPT SMALL LETTER D;Lm;0;L;<sub> 0064;;;;N;;;;;
1DFF2;LATIN SUBSCRIPT SMALL LETTER D WITH STROKE;Lm;0;L;<sub> 0111;;;;N;;;;;
1DFF3;LATIN SUBSCRIPT SMALL LETTER OPEN E;Lm;0;L;<sub> 025B;;;;N;;;;;
1DFF4;LATIN SUBSCRIPT SMALL LETTER F;Lm;0;L;<sub> 0066;;;;N;;;;;
1DFF5;LATIN SUBSCRIPT SMALL LETTER G;Lm;0;L;<sub> 0067;;;;N;;;;;
1DFF6;LATIN SUBSCRIPT SMALL LETTER GAMMA;Lm;0;L;<sub> 0263;;;;N;;;;;
1DFF7;LATIN SUBSCRIPT SMALL LETTER PHI;Lm;0;L;<sub> 0278;;;;N;;;;;
1DFF8;LATIN SUBSCRIPT SMALL LETTER TURNED R;Lm;0;L;<sub> 0279;;;;N;;;;;
1DFF9;LATIN SUBSCRIPT SMALL LETTER R WITH TAIL;Lm;0;L;<sub> 027D;;;;N;;;;;
1DFFA;LATIN SUBSCRIPT SMALL CAPITAL R;Lm;0;L;<sub> 0280;;;;N;;;;;
1DFFB;LATIN SUBSCRIPT SMALL CAPITAL INVERTED R;Lm;0;L;<sub> 0281;;;;N;;;;;
1DFFC;LATIN SUBSCRIPT SMALL LETTER W;Lm;0;L;<sub> 0077;;;;N;;;;;
1DFFD;LATIN SUBSCRIPT SMALL LETTER Y;Lm;0;L;<sub> 0079;;;;N;;;;;
1DFFE;LATIN SUBSCRIPT SMALL LETTER Z;Lm;0;L;<sub> 007A;;;;N;;;;;
1DFFF;LATIN SUBSCRIPT SMALL LETTER EZH;Lm;0;L;<sub> 0292;;;;N;;;;;

References

American Anthropological Association (1916) *Phonetic Transcription Of Indian Languages*.

Andersen (1977) "Dansk Dialektologi i Tiåret 1965–74," *Danske Studieren*.

Juan Lope Blanch, ed. (1990) *Atlas lingüístico de México*. Vol. I, Fonética. Centro de Estudios Lingüísticos y Literarios, El Colegio de México, Fondo de Cultura Económica.

Bickford & Floyd (2006) *Articulatory Phonetics*, 4th edition. SIL International.

Boas, Goddard, Sapir & Kroeber (1916) *Phonetic Transcription of Indian Languages: Report of Committee of American Anthropological Association* (Smithsonian Miscellaneous Collections 66.6)

Frasz Boas (1892) Notes on the Chemakum Language. *American Anthropologist* 5(1): 37-44.

Donald Burquest, trans. Giuliana López Torres (2009) *Análisis fonológico*. SIL International.

Pedro Martín Butragueño (2014) *Fonología variable del español de México*. Vol. I Procesos segmentales.

Charbonneau & Jacques (1972) "[t_s] et [d_z] en français canadien," in Valdman (ed.) *Papers in Linguistics and Phonetics to the Memory of Pierre Delattre*.

Jingjing Fan (2018) 'Newly Identified Khotanese Fragments in the "Bodhisattva Compendium" and Their Chinese, Pāli and Sanskrit Parallels'. In the *Annual Report of the International Research Institute for Advanced Buddhology (ARIRIAB) at Soka University for the Academic Year 2017* (vol. XXI).

Nina Grønnum (2005) *Fonetik og Fonologi: Almen og dansk*. Akademisk forlag, Copenhagen.

Louis Leonor Hammerich (1934) *Tysk fonetik*

Hickey (2014) *The Sound Structure of Modern Irish*.

Kalnyn' (1973): Л.Э. Калнынь, *Опыт моделирования системы украинского диалектного языка*. Наука.

Kalnyn' & Maslennikova (1981): Л.Э. Калнынь & Л.И. Масленникова, *Сопоставительная модель фонологической системы славянских диалектов*. Наука, Moscow.

Patricia Keating (2018) 'Some challenges for the IPA from languages of China'. 13th Phonetic Conference of China, Jinan University, Guangzhou.

Ling Feng (2007) 'The articulatory and acoustic study of fricative vowels in Suzhou Chinese'. *Proceedings of ICPhS XVI*, International Congress of Phonetic Sciences.

Pulleyblank (1962) "The Consonantal System of Old Chinese," *Asia Major* 9: 2.

RFE (1915) 'Alfabeto fonético de la revista de filología española', *Revista de filología española*, vol. II.

Skjærvø (2002) *Khotanese Manuscripts from Chinese Turkestan in the British Library*, part 2, vol. 5.

Wickelgren, Wayne (1969) 'Context sensitive coding, associative memory, and serial order in (speech) behavior'. *Psychological Review*, 76 (1).

Figures

Subscript b (ᵦ)

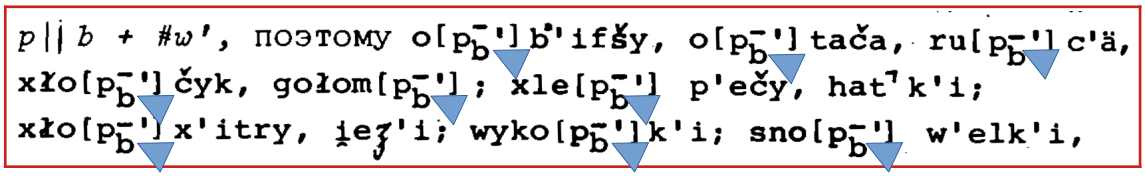


Figure 4. Kalnyn' & Maslennikova (1981: 350). The archiphoneme is the result of the neutralization of voicing (*p* vs *b*) and of palatalization (*p*¯ vs *p'*).

Subscript c (ᶜ)

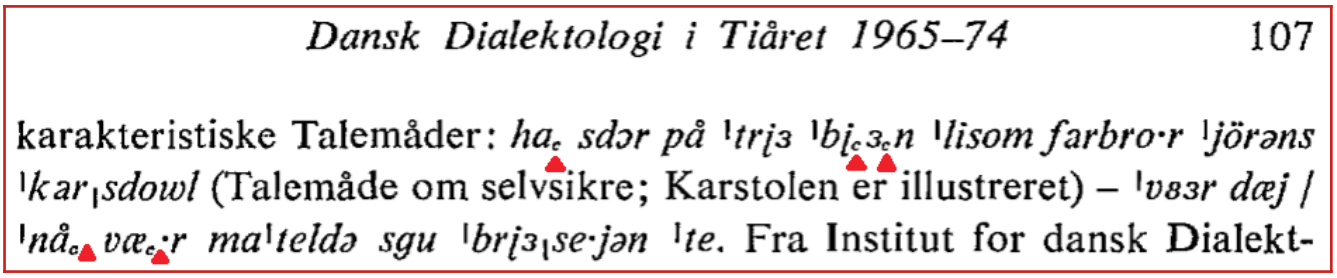


Figure 5. Andersen (1977:107). This 'nasal curl', semantically equivalent to an ogonek, is typeset in the Dania font that has been shared with me as a subscript 'c', and this published example appears to be a subscript 'c' as well.

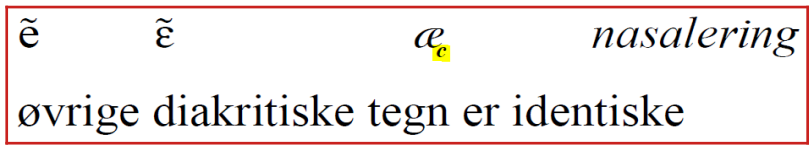
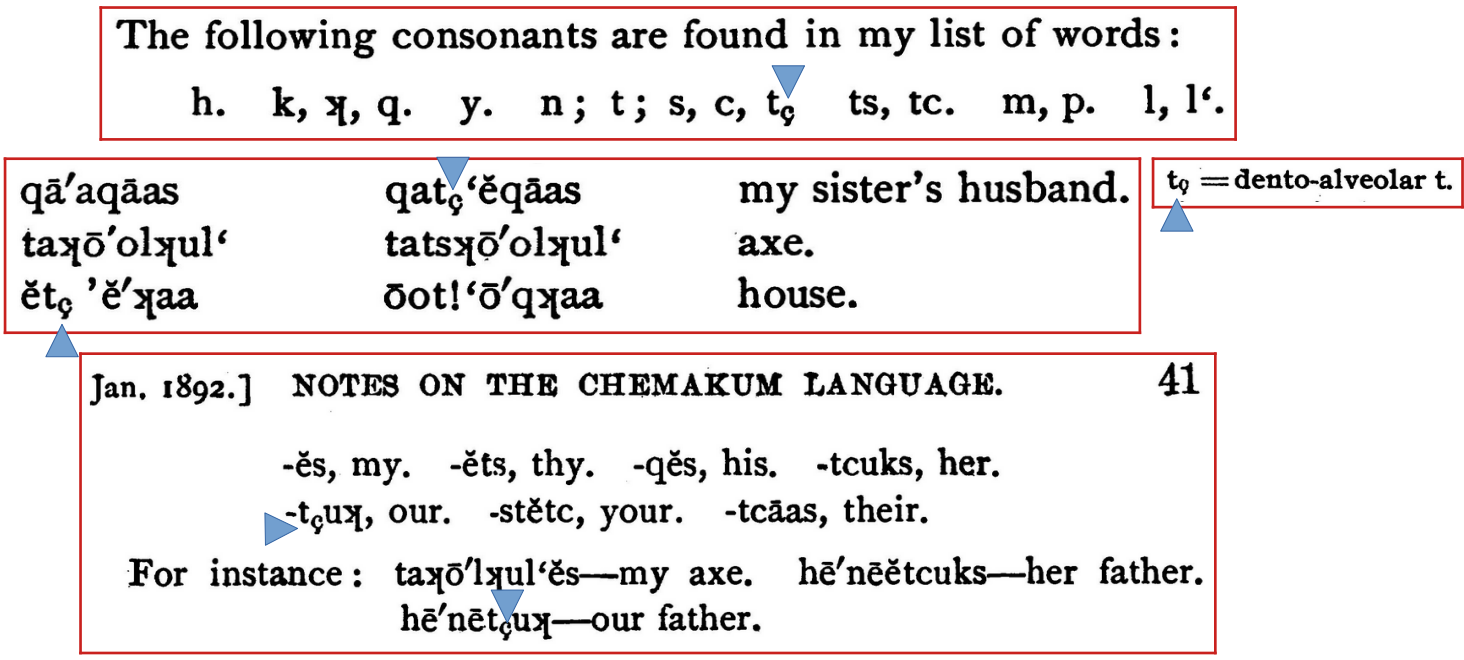


Figure 6. Grønnum (2005: 419). In this publication as well, the 'nasal curl' is typeset as a subscript 'c'.


The following consonants are found in my list of words:

h. k, ɬ, q. y. n; t; s, c, $t_ç$ ts, tc. m, p. l, l‘.

qā′aqāas	qat$_ç$‘ĕqāas	my sister’s husband.
takɬō′olkul‘	tatskɬō′olkul‘	axe.
ĕt$_ç$ ’ĕ′ɬaa	ōot!‘ō′qɬaa	house.

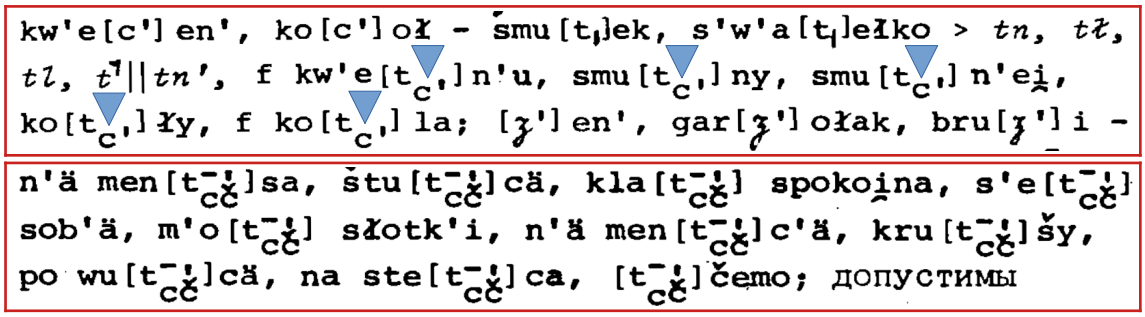
$t_ç$ = dento-alveolar t.

Jan. 1892.] NOTES ON THE CHEMAKUM LANGUAGE. 41

-ĕs, my. -ĕts, thy. -qĕs, his. -tcuks, her.
-t$_ç$uk, our. -stĕtc, your. -tcāas, their.

For instance: takɬō′lkul‘ĕs—my axe. hē′nēĕtcuks—her father.
hē′nēt$_ç$uk—our father.

Figure 7. Boas (1892: 37 fn, 38, 39, 41). As with superscript ⟨ᶜ⟩, the subscript should be rendered as a ⟨$_c$⟩ plus combining cedilla.


kw'e[c'] en', ko[c'] oɬ - šmu[t$_ç$]ek, s'w'a[t$_ç$]eɬko > *tn*, *tɬ*, *tl*, *t*^{n}||*tn'*, f kw'e[t$_ç$,] n'u, smu[t$_ç$,] ny, smu[t$_ç$,] n'ei, ko[t$_ç$,] ɣy, f ko[t$_ç$,] la; [ʒ'] en', gar[ʒ'] oɬak, bru[ʒ'] i -

n'ä men[t$_{cč}$]sa, štu[t$_{cč}$]cä, kla[t$_{cč}$] spokoina, s'e[t$_{cč}$] sob'ä, m'o[t$_{cč}$] sɬotk'i, n'ä men[t$_{cč}$]c'ä, kru[t$_{cč}$]šy, po wu[t$_{cč}$]cä, na ste[t$_{cč}$]ca, [t$_{cč}$]čemo; допустимы

Figure 8. Kalnyn’ & Maslennikova (1981: 360, 366). ⟨t$_{c}^{\prime}$⟩ [with a combining apostrophe on the subscript modifier *c* that does not align properly in the font used for this proposal] is an archiphoneme that conflates /t/ and /c’/ (IPA [t͡sʲ]). Subscript *č* should be handled with a combining diacritic.

Subscript d ($_d$)

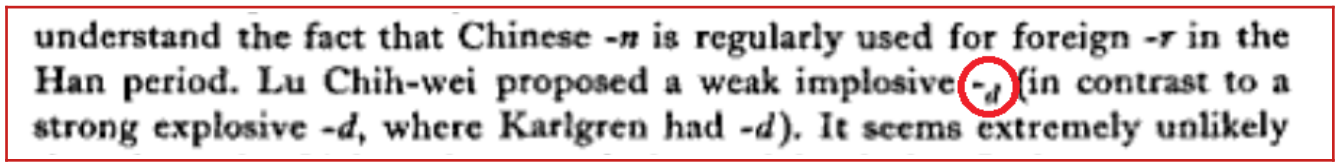

understand the fact that Chinese *-n* is regularly used for foreign *-r* in the Han period. Lu Chih-wei proposed a weak implosive *-$_d$* (in contrast to a strong explosive *-d*, where Karlgren had *-d*). It seems extremely unlikely

Figure 9. Pulleyblank (1962: 215). Later on the page Pulleyblank says that all three final plosives appear to have been implosive. (“Implosive” here seems to mean unreleased.) In this notation the three would be $^{-}_{b}$, $^{-}_{d}$, $^{-}_{g}$.

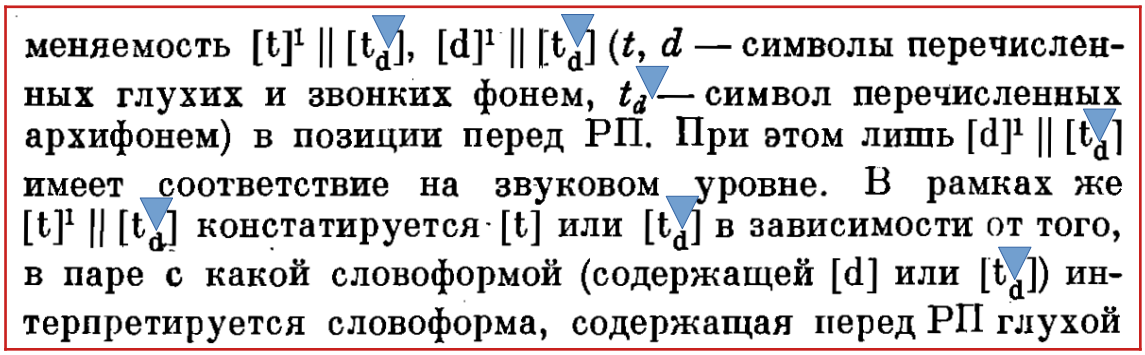
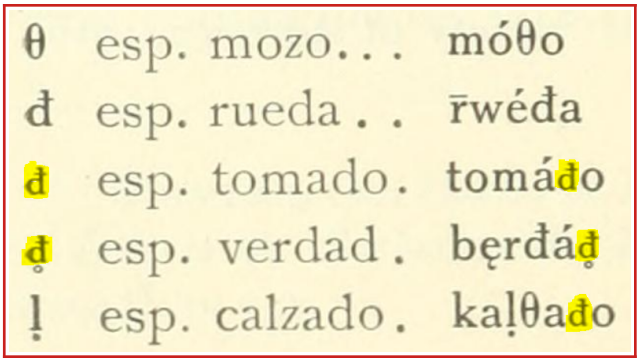

меняемость [t]¹ || [t_d], [d]¹ || [t_d] (*t*, *d* — символы перечисленных глухих и звонких фонем, t_d — символ перечисленных архифонем) в позиции перед РП. При этом лишь [d]¹ || [t_d] имеет соответствие на звуковом уровне. В рамках же [t]¹ || [t_d] констатируется [t] или [t_d] в зависимости от того, в паре с какой словоформой (содержащей [d] или [t_d]) интерпретируется словоформа, содержащая перед РП глухой

Figure 10. Kalnyn' (1973: 327). An sample of the neutralization of a voicing contrast is given in IPA: [t] and [d] are an example of a voicing pair; both neutralize to [t_d]. (The specific illustrations for Ukrainian are presented in Cyrillic.)

Subscript d with stroke (ᵭ)

The traditional Spanish phonetic alphabet (the RFE, or *Revista de filología española*, alphabet) uses reduced letters, equivalent to subscripts, for reduced sounds. A reduced ⟨đ⟩ is specifically provided for. The RFE alphabet is still used and taught today in Spanish-language universities, notably in Spain and Mexico. Apart from the reduced sounds, the RFE is covered by Unicode.


θ esp. mozo. . . móθo
đ esp. rueda . . r̄wéđa
đ esp. tomado. tomáđo
đ̥ esp. verdad. bęrđáđ̥
ļ esp. calzado. kaļθađo

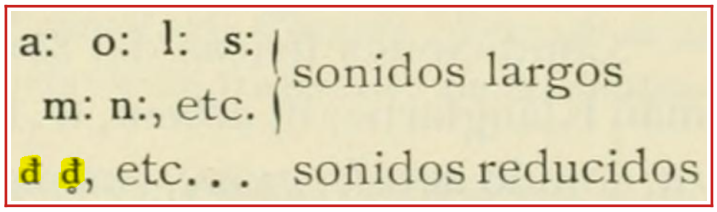

a: o: l: s: m: n:, etc. } sonidos largos
đ đ̥, etc. . . sonidos reducidos

Figure 11. RFE (1915: 374–375). Contrast between normal ⟨đ⟩, small ⟨ᵭ⟩ and small voiceless ⟨đ̥⟩, in the consonant table. The RFE alphabet chart has been reproduced many times over the past century, including in 2020 by Alexander Iribar of the phonetics laboratory at the University of Deusto in the Spanish Basque country. The second clip shows that reducing font size to indicate phonetic reduction is a productive convention, but I've only seen ⟨đ⟩ used in this way.

Subscript open e (ɛ)

фонемы <e$_{ɛ}$> и <i$_{e}$> могут оформлять одну и ту же словоформу: ро́бй<e$_{ɛ}$>т (фонетически *ро́бйет*) и ро́бй<i$_{e}$>т (фонетически *ро́бйіт*). Словоформы, в которых <e$_{ɛ}$> варьируется с <e>, могут содержать и <i$_{e}$>, т.е. (<e>^{1}||<e$_{ɛ}$>)^{1}||-

[e] - [ɛ]: с[ɛ]ла (им. мн.) - с[e]р > с<e$_{ɛ}$>ло́, с<e$_{ɛ}$>ро́ватка (фонетически *сɛ́ла, сер > село́, серо́ватка*), т.е.

<i> <y> с<i>лне́ц'е - с<y>ка́ти - с<e$_{ɛ}$>ро́ватка, к'е-
<e$_{ɛ}$> <o> с<e$_{ɛ}$>ле́ц'е - с<o>к'е́ра - с<a>моно́нка

Figure 12. Kalnyn' & Maslennikova (1981: 155, 158). The distinction between the letters *e* and *ɛ* is Greek/IPA, filling in a gap in Cyrillic notation.

Subscript f (f)

поэтому po[w$_{f}^{-}'$] b'iwana;[w$_{f}^{-}'$]f'ilm'ä, ʒ'en', o[w$_{f}^{-}'$] sy, štra[w$_{f}^{-}'$] w'elk'i, šča[w$_{f}^{-}'$] z'elony; kre[w$_{f}^{-}'$] gęnsta, le
i̯ego; ro[w$_{f}^{-}'$] k'i; [z]ydel, [z] agrafko, zaw'o[z]a,

Figure 13. Kalnyn' & Maslennikova (1981: 350). In these examples the letters *f* and *w* have their Polish Latin values of [f] and [v] rather than the values [ɸ] and [β] they take in Cyrillic phonetic notation.

Subscript g (g)

k||*k'* # *x'*, *k*||*k'*||*g*||*g'* + #*i*, поэтому i̯a[k$_{g}^{-}'$]by, [k$_{g}^{-}'$]p'iny, ʒ'e[k$_{g}^{-}'$]c'u, ro[k$_{g}^{-}'$] žäš; i̯a[k$_{g}^{-}'$]p'iwo, rydlufka, żyto, hycel, g'ili, i̯äxali, wypušča; to[k$_{g}^{-}'$] z'b'ity, w'il[k$_{g}^{-}'$]x'itry; перед #*V* допустимо [b$^{-}'$]^{1}||-

[k$_{r}^{-}'$]/[k$_{g}^{-}'$]

Figure 14. Kalnyn' & Maslennikova (1981: 350, 372). The archiphoneme is the result of the neutralization of voicing (*k* vs *g*) and of palatalization (*k*$^{-}$ vs *k'*). The authors use an IPA-like single-loop 'g' in most cases (as in the upper illustration above), though the transcription is not IPA (e.g. *c*, *w* above have their Polish values of [v, t͡s]). Occasionally however they use a two-loop 'g' (lower illustration). Because the usual convention is to use a standard Latin 'g' and because there is no contrast, I

request the standard ‘g’ in this proposal.

Subscript Latin gamma (ᵧ)

Used for simultaneous velarization as opposed to a velar offglide (Bickford & Floyd) or relatively weak velarization (Hickey).

In the Irish of Roscommon/East Galway[25], Ring and Cape Clear voiced sibilants have been reported as the outcome of nasalising /s/, e.g. *i Sasana* [ɪ zasənᵧə] ‘in England’.

Figure 15. Hickey (2011) *The Dialects of Irish*, p. 31. Subscript notation is described next.

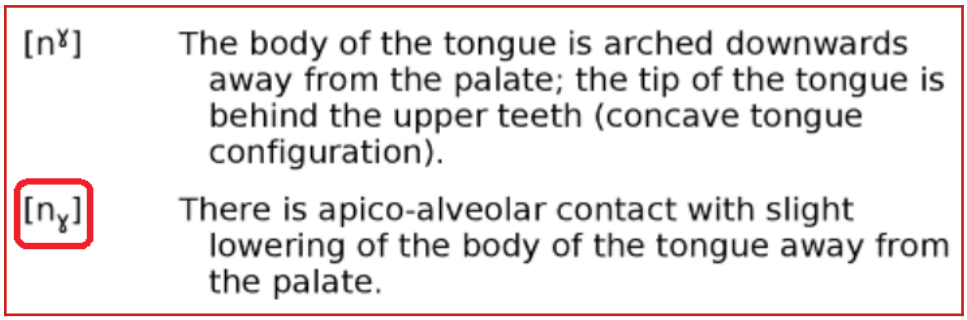


Figure 16. Hickey (2014: list just before §1.8.5; no page number in ebook). In the paragraph before §1.2 Hickey explains that a subscript gamma is used for southern dialects that have weak velarization. There is a similar distinction between [nʲ] (palatal) and [nⱼ] (palatalized alveolar).

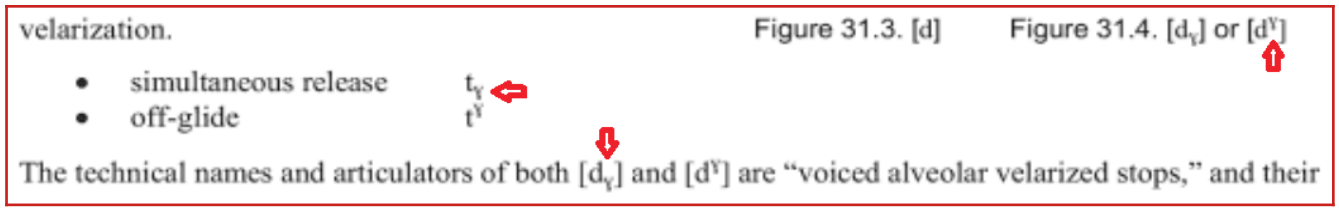


Figure 17. Bickford & Floyd (2006: 162)

Subscript ɹ ɽ ʀ ʁ (ᵢ ᵣ ᵣ ᵦ)

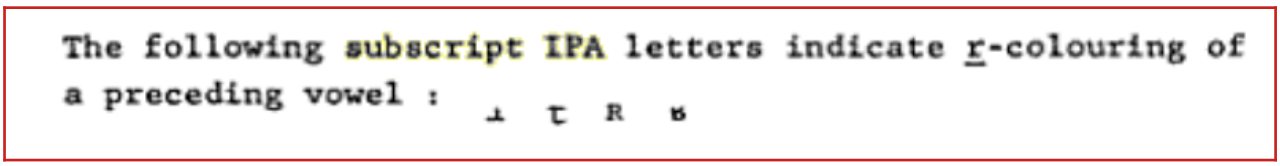


Figure 18. Penhallurick (1991) *The Anglo-Welsh Dialects of North Wales*, p. xviii. Subscript ɹ ɽ ʀ ʁ are used in the data.

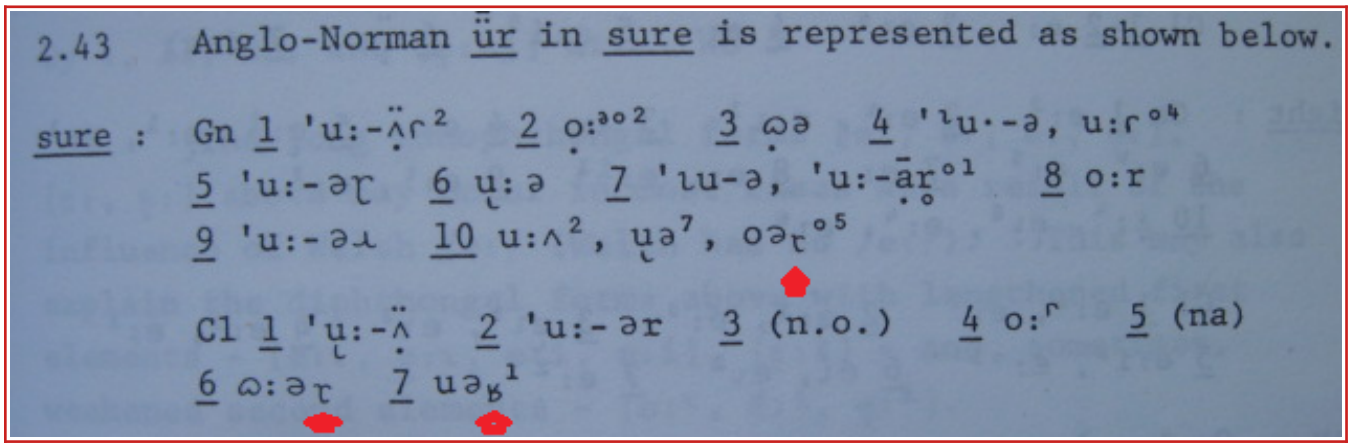


Figure 19. Penhallurick (1991: 85). Subscript ɽ ʁ along with superscript ʳ.

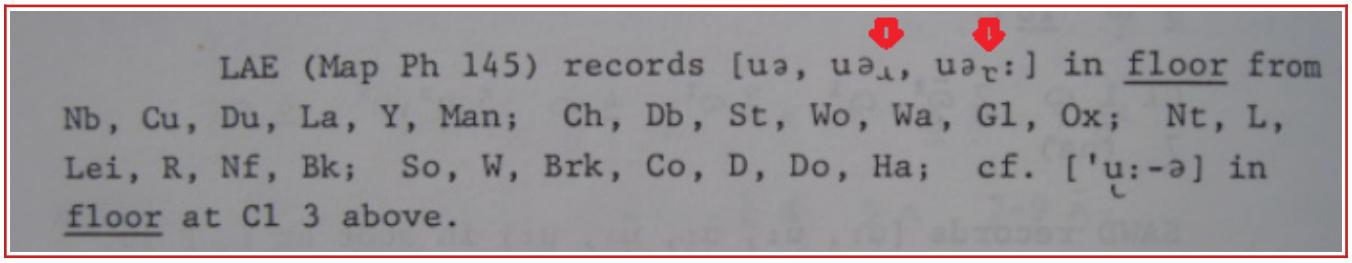


Figure 20. Penhallurick (1991: 85). Subscript ɹ ɽ. I have not found subscript ʁ in this volume, but ʁ is used in the data it was drawn from (see first figure above). The data is published in volume 3 of D.R. Parry (director, U. of Swansea) & Penhallurick (ed.) *Survey of Anglo-Welsh Dialects*, which I don't have access to.

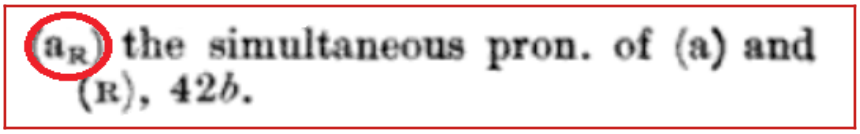


Figure 21. Subscript ʁ in Ellis (1889) *On Early English Pronunciation*, part V, p. 78.

Subscript w ($_w$)

Como un segundo ejemplo de labialización obsérvense los datos del angas en (161). En estos datos, el símbolo [$_w$] pequeño debajo de ciertas consonantes representa el redondeamiento de los labios durante la articulación de la consonante sin una fuerte labialización en la distensión de ésta. La letra sobre escrita [^{w}] es una distensión labializada normal de la consonante:

(161)	/po/	[p$_w$o]	'boca'
	/bum/	[b$_w$um]	'gorro'
	/tu/	[t$_w$u]	'matar'
	/du/	[d$_w$u]	'oler'
	/ko/	[k^{w}o]	'o'

Figure 22. Burquest (2009: 130)

Labialized consonants, that is, consonants pronounced with simultaneous lip-rounding, are to be indicated by means of inferior *w* closely following the character. Thus, l_w indicates an *l* pronounced with markedly rounded lips; similarly, k_w indicates a *k* with simultaneous lip-rounding (not to be confused, of course, with *kw*).

Figure 23. Boas, Goddard, Sapir & Kroeber (1916: 15)

Subscript y ($_y$)

6 [ṣḍe ysaṃ]tha kuśalamūla hataḍarāṃjs$_y$äya ni īndä ṣä kiḍe (du)ṣkaru ka ye tt$_y$ānu śiru yuḍu īndä.

Figure 24. Fan (2018: 329), reproducing Skjærvø (2002: 343).

Palatalized consonants, that is, consonants modified by the simultaneous articulation of a large part of the surface of the tongue against the palate (in other words, by the tongue taking *y*-position), are to be indicated by closely following inferior *y*. Thus, n_y indicates a palatalized dental *n*. The ordinary so-called "palatal" *l* and *n* are probably best considered as palatalized dorsal *l* and *n* and should thus, strictly speaking, be indicated by λ_y (Italian *gl*) and ν_y (Italian *gn*); l_y and n_y would, however, be the normal methods of representing these consonants.

Figure 25. Boas, Goddard, Sapir & Kroeber (1916: 15)

Subscript z ($_z$)

Subscripts are used for fricative release in some French sources, filling the role played by superscripts in English. An example is subscript s and z for the frication of [t^s] and [d^z] (allophones of /t/ and /d/ in Quebequois French).

Patricia Keating used it during her tenure as president of the IPA.

conclure globalement qu'il ne s'agissait pas d'une véritable palatalisation, mais d'un assibilation. Deux affriquées, une sourde [t_s] et une sonore [d_z], apparaissaient a contact de [i] et de [y] dans une même syllabe. Ces consonnes affriquées avaien

Figure 26. Charbonneau & Jacques (1972: 77). Subscript *z* is also used in the title of the article, "[t_s] et [d_z] en français canadien."

devant [y], dans la phrase: 'il y a du vent' [ilja$d_z$yvɑ̃].

Figure 27. Charbonneau & Jacques (1972: 87)

Ling (2007) shows narrower and fronter constriction for fricative vowel [i_z] compared to [i]:

igure 2: Palatograms and linguagrams of [i_z] and [i]
f a male speaker.

Figure 28. Keating (2018: 27).

There are totally 12 vowels in Suzhou Chinese, which are [i_z y_z u i y ø ɛ o æ ɑ ɿ ʯ].Two pairs of

vowel quality, the test words associated with high level tone [44] with zero initial consonant were selected. The test words were:
[i_z^{44}] (coat) [i^{44}] (smoke).

Figure 29. Ling (2007: 573).

Figure 30. Kalnyn' & Maslennikova (1981: 337, 378). The hacek should be handled with a combining diacritic.

Subscript з (з)

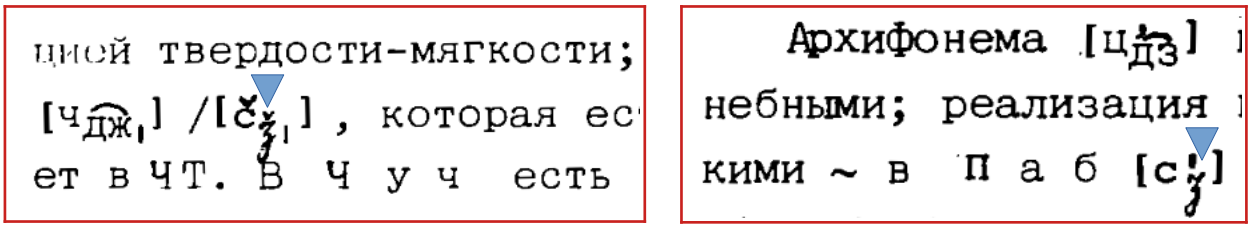


Figure 31. Kalnyn' & Maslennikova (1981: 338, 378). In the typical romanization of the countries of the ex–Soviet Union, ⟨c з⟩ are affricates [ts, dz] (Cyrillic *ц дз*), and ⟨č ǯ⟩ are [tʃ dʒ] (Cyrillic *ч дж*).

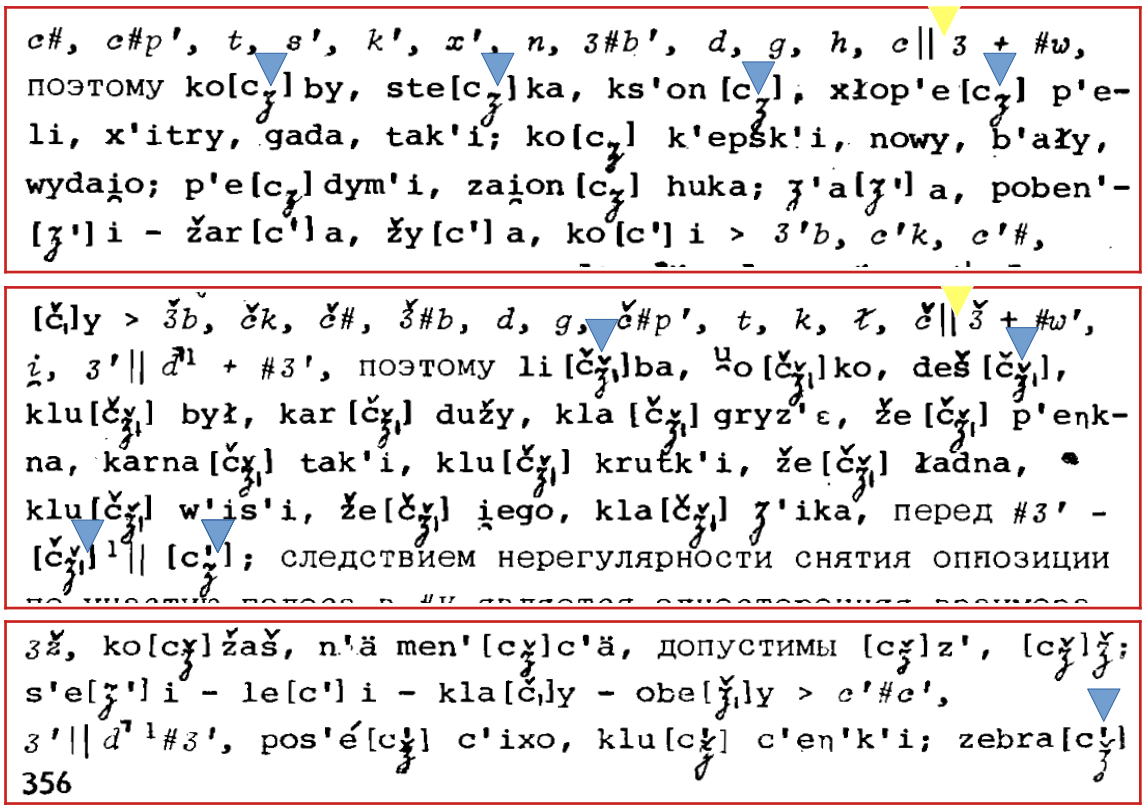


Figure 32. Kalnyn' & Maslennikova (1981: 344, 356). The c-з and č-ǯ contrasts are presented in the first lines (marked in yellow), with $c_з$ and $č_з$ archiphonemes illustrated in various words (examples marked in blue). The $c^{v}_{з}$ archiphoneme is a

neutralization of both voicing and palatalization: {ts, dz, tʃ dʒ}.

Subscript Latin phi (ᵩ)

[ɸ̬]	sonorizada
[ᵩᶠ]	bilabiodental
[ɸᶠ]	bilabiodental con predominio del elemento bilabial
[fᵠ]	bilabiodental con predominio del elemento labiodental
[ᵩˣ]	labiovelar con predominio del elemento labial
[ₓᵠ]	labiovelar con predominio del elemento velar
[ᵩʰ]	labiofaríngea

[ᵩʰ] labiofaríngea

Figure 33. Butragueño (2014: 29 *ff*): IPA equivalents of the RFE Spanish phonetic alphabet, as used e.g. in the *Atlas Lingüístico de México*, contrasting baseline, superscript and subscript *ɸ*.

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:** *Additional phonetic click letters*
2. Requester's name: *Kirk Miller*
3. Requester type (Member body/Liaison/Individual contribution): *individual*
4. Submission date: *2021 August 16*
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-G*
2. Number of characters in proposal: *17*
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)? *no*

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? *Author is a member 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) *transcription*
   Reference:
5. Are the proposed characters in current use by the user community? *yes*
   If YES, where? Reference:
6. After giving due considerations to the principles in the P&P document must the proposed characters be entirely in the BMP? *no*
   If YES, is a rationale provided?
   If YES, reference:
7. Should the proposed characters be kept together in a contiguous range (rather than being scattered)? *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? *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: