L2/24-219

Unicode request for subscript *w y z* and ɣ

Kirk Miller, kirkmiller, gmail.com 2024 October 18

Semantically distinct subscripts are used in several phonetic traditions. It would seem that the distinction between superscript and subscript is at least partly iconic: Superscript modifiers such as IPA ⟨kʷ dʲ⟩ have the visual appearance of being off-glides, whereas the old-style IPA combining subscripts ⟨k̟ d̟⟩ do not; spacing-subscript ⟨kw dj⟩ are found as graphic approximations of the latter.

Some phoneticians use subscript modifier letters for simultaneous articulation and restrict superscript modifiers to sequential sounds such as onsets, on- and off-glides, and releases; others use subscripts for shades of sound and superscripts for full co-articulation. For instance, Hickey (2014) uses subscript ɣ and *j* for light velarization and palatalization, contrasting with superscript ɣ and *j* for full co-articulation, while Bickford & Floyd (2006) use subscript ɣ for simultaneous velarization and superscript ɣ for a velar offglide. Burquest (2009) does the same for *w*, a tradition dating back to Boas et al. (1916), who in the Americanist tradition used *y* for *j*. Penhallurick (1991) uses subscript rhotic letters such as ⟨aɹ⟩ for r-colored vowels, and superscripts such as ⟨ʴ⟩ for weakly articulated but still consonantal rhotic sounds. Keating (2018) and Ling (2007) use subscript *z* to mark a fricative vowel, whereas a superscript *z* would indicate consonantal frication; Keating used this convention during her tenure as president of the IPA. In the French tradition subscripts rather than superscripts are used for fricative release, such as the allophones [tₛ] and [dᵤ] of Quebequois French.

Characters

Subscripts for phonetic transcription

ʷ 209D LATIN SUBSCRIPT SMALL LETTER W. Figure 1 *ff.*

ʸ 209E LATIN SUBSCRIPT SMALL LETTER Y. Figure 4, Figure 5 *ff.*

ᶻ 209F LATIN SUBSCRIPT SMALL LETTER Z. Figure 10 *ff.*

ᵞ 1DFD0 LATIN SUBSCRIPT SMALL LETTER GAMMA. Figure 15 *ff.*

Properties

209D;LATIN SUBSCRIPT SMALL LETTER W;Lm;0;L;<sub> 0077;;;;N;;;;;
209E;LATIN SUBSCRIPT SMALL LETTER Y;Lm;0;L;<sub> 0079;;;;N;;;;;
209F;LATIN SUBSCRIPT SMALL LETTER Z;Lm;0;L;<sub> 007A;;;;N;;;;;

1DFD0;LATIN SUBSCRIPT SMALL LETTER GAMMA;Lm;0;L;<sub> 0263;;;;N;;;;;

Charts

There is space for subscript *x y z* in the Superscripts and Subscripts block in the BMP.

Superscripts and Subscripts

2070 209F

	207	208	209
0	0	0	a
1	i	1	e
2		2	o
3		3	x
4	4	4	ə
5	5	5	h
6	6	6	k
7	7	7	l
8	8	8	m
9	9	9	n
A	+	+	p
B	−	−	s
C	=	=	t
D	(	(	w
E	)	)	y
F	n		z

The ɣ should be placed in Latin Extended-G.

Latin Extended-G

1DF00 **1DFFF**

	1DF0	1DF1	1DF2	1DF3	1DF4	1DF5	1DF6	1DF7	1DF8	1DF9	1DFA	1DFB	1DFC	1DFD	1DFE	1DFF
0	fŋ	ĸ	ʤ	ɢ										ɣ	ʤ	ɭ
1	ϱ	ł	ʤ	ɣ											ʤ	ɳ
2	ɔ	ʤ	tɬ	ħ											ʃ	ƫ
3	ʞ	ʄ	tɬ	ɸ											ʃ	ψ
4	ɬ	ŋ	tθ	q											ʧ	ω
5	ɮ	ɻ	ɗ	ʀ											ʧ	ɖ
6	ꭗ	ʃ	ɭ	ĸ											ʒ	ɦ
7	ɑ	ʧ	ɳ	ɾ											ʒ	ŋ
8	ɭ	ʒ	ɣ	ʦ										ɥ	ᴅ	ʂ
9	ʄ	ʤ	s	y										ɟ	ᴀ	ʐ
A	ǃ	ɨ	ƭ	ʃ										ᵿ	ᴇ	ɕ
B	ʄ	ϱ	ʣ	β										w	ı	ƙ
C	ʄ	ʧ	ʧ	θ										η	ɩ	ɓ
D	ʓ	ç	dʃ	χ										ɼ	ɥ	ɖ
E	ʆ	ʂ	ʣ	ɥ										ȸ	ɥ	ƭ
F	ɕ	dð	ð	w										ȹ	ɗ	ʃ

References

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)

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

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*.

Hickey (2011) *The Dialects of Irish*.

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

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.

Uli Kozok (2015) *A 14th Century Malay Code of Laws*. ISEAS, Singapore.

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

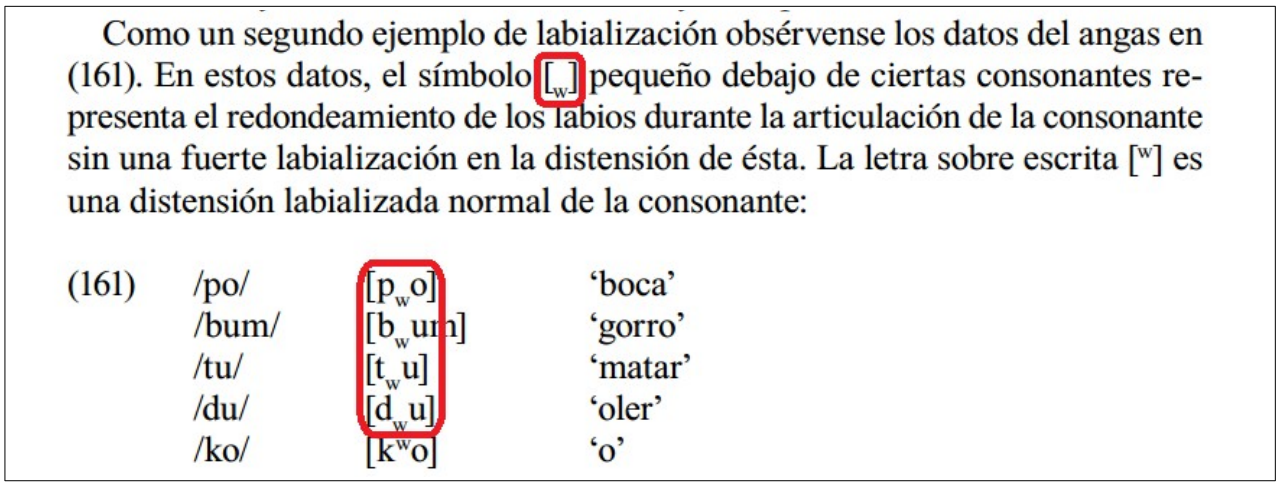
Luo & Wang (2002): 罗常培、王均《普通语音学纲要》商务印书馆.

Penhallurick (1991) *The Anglo-Welsh Dialects of North Wales*.

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

George Trager (1974) Writing and Writing Systems. In Sebeok, ed., *Current Trends in Linguistics*, vol. 12, *Linguistics and Adjacent Arts and Sciences*. Mouton.

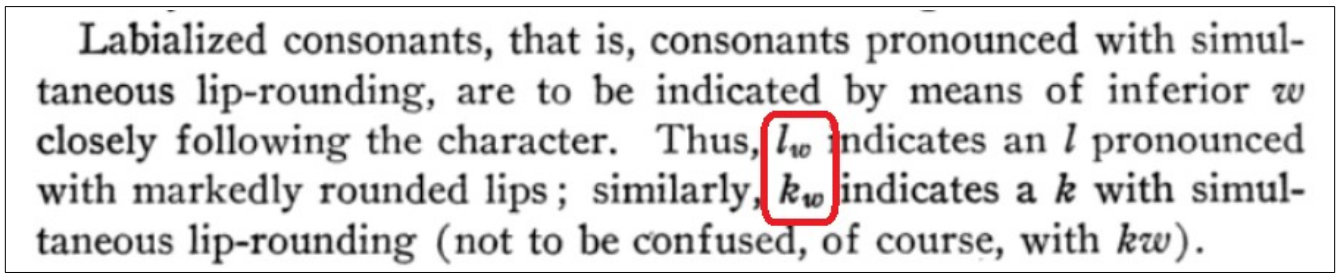
Figures



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:

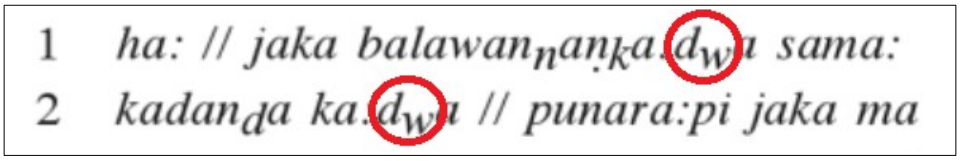
(161)	/po/	[p$_w$o]	'boca'
	/bum/	[b$_w$um]	'gorro'
	/tu/	[t$_w$u]	'matar'
	/du/	[d$_w$u]	'oler'
	/ko/	[k^{w}o]	'o'

Figure 1. 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 2. Boas, Goddard, Sapir & Kroeber (1916: 15).



1 *ha: // jaka balawan$_n$an$_k$a.d$_w$u sama:*
2 *kadan$_d$a ka.d$_w$u // punara:pi jaka ma*

Figure 3. Kozok (2015: 71). May be twice subscribed for a double conjunct (e.g. p. 72, line 26/4), but that is not semantically distinct from a simple conjunct and so does not require additional Unicode support.

some instances of 'vowel length', and which had various allophones before different vowels — perhaps a glottal stop, [ʔ], before /a/; a palatalized glottal stop, [ʔ$_y$] with voiceless palatal glide from [ʔ] to the vowel, before /e/; some velar glide, [ʔ$_x$], before /o/; labialization before /u/, [ʔ$_w$]; before /i/ the palatalization was stronger than before /e/. An initial syllable /'a/ would be written, in Semitic fashion, by the symbol

Figure 4. Trager (1974: 402). Subscript *w*, *x* and *y*. Trager contrasts subscripting with supercripting elsewhere on the page.

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

Figure 5. 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 6. Boas, Goddard, Sapir & Kroeber (1916: 15)

5 / uysnaurāṇu śir^{e}atete yanāma kiḍe duṣkara. tta kāma drrai padya cu ttānu uysnaurāṇu haṃjs^{i}a-
6 [ṣḍe ysaṃ]tha kuśalamūla hataḍarāṃjs$_y$äya ni Indä ṣä kiḍe (du)ṣkaru ka ye tt$_y$ānu śiru yuḍu Indä. ttina cu
7 uysnau]rāṇu ṣe' paḍā. kye vātcā anyattīrthya o vā paḍā añattirthiya väta o^{au} nu vātcu ttirthānu duiṣta. tta
8 / x šširu g$_{yu}$u g$_{yi}$ndi tta ṣā vātcu šā'ta^{a} balysūñavūysai kiḍe duṣkara :ll : ll .

Figure 7. Skjærvø (2002: 343).

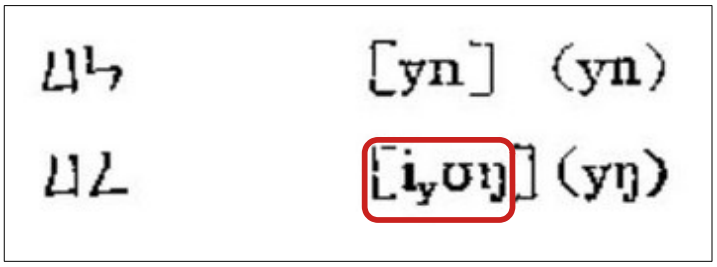


Figure 8. Luo & Wang (2002: 23). Transcriptions of the *zhuyin* letter 〈ㄩ〉.

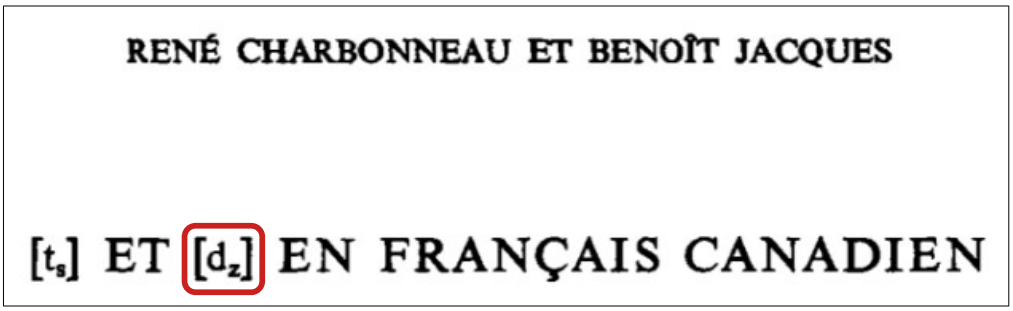


Figure 9. Charbonneau & Jacques (1972: 77). Subscript *z* used in the title of the article.

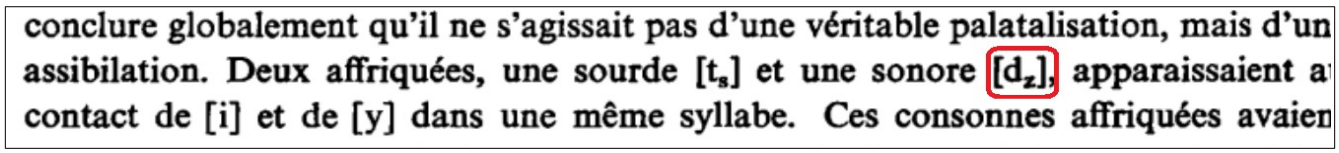

conclure globalement qu'il ne s'agissait pas d'une véritable palatalisation, mais d'un assibilation. Deux affriquées, une sourde [ts] et une sonore [dz], apparaissaient a contact de [i] et de [y] dans une même syllabe. Ces consonnes affriquées avaien

Figure 10. Charbonneau & Jacques (1972: 77).

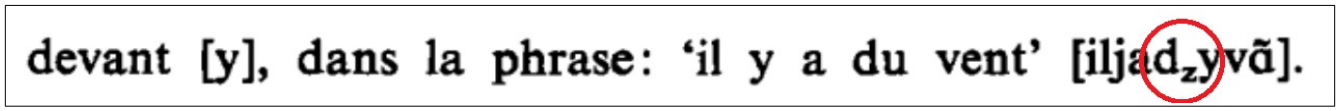
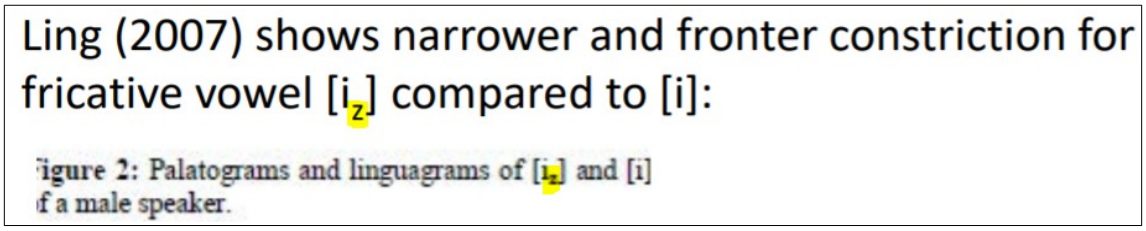

devant [y], dans la phrase: 'il y a du vent' [iljadzyvã].

Figure 11. Charbonneau & Jacques (1972: 87)


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

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

Figure 12. Keating (2018: 27).

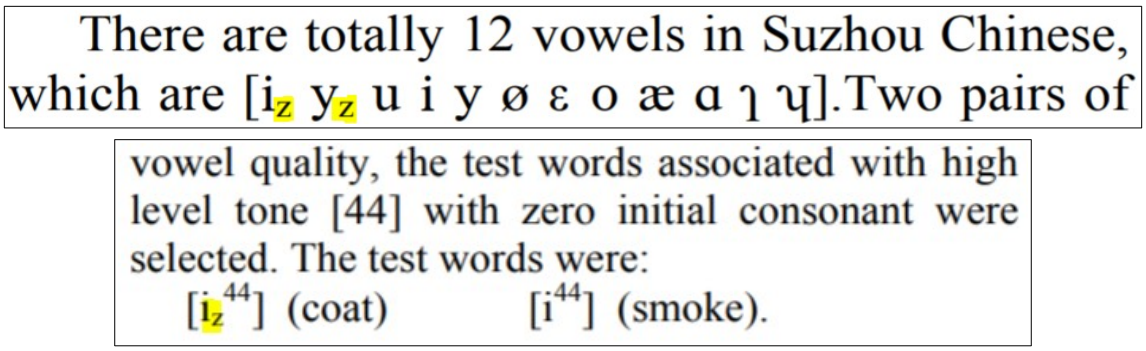


Figure 13. Ling (2007: 573).

Figure 14. Kalnyn' & Maslennikova (1981: 337, 378). The hacek should be handled with a combining diacritic.

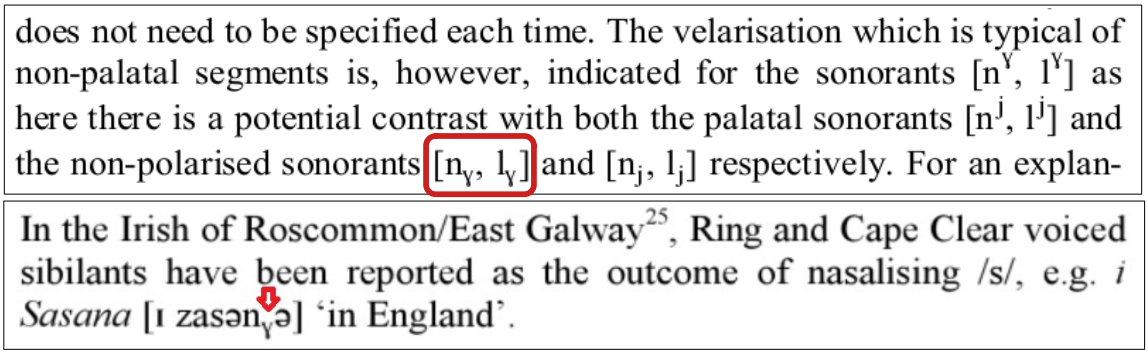


Figure 15. Hickey (2011: 30, 31). Subscript notation is described below.

Table 57. Polarity cline for *n*- and *l*-sounds

maximal palatality				maximal velarity
n^j, l^j		n_j, l_j $n_ɣ$, $l_ɣ$		$n^ɣ$, $l^ɣ$
high polarity	—	low polarity	—	high polarity

It is clear that the low-polarity items contrast with the high polarity items and this contrast is systemic in many dialects, e.g. *coll* /kʌ$l^ɣ$/ 'hazel' # *col* /kʌ$l_ɣ$/ 'aversion; incest'. But quite another issue is whether the low-polarity

[$n_ɣ$] There is apico-alveolar contact with slight lowering of the body of the tongue away from the palate.

Table 3. Fourway distinction for *n*- and *l*-sounds

l^j	l_j	$l_ɣ$	$l^ɣ$
buille 'blow'	*buile* 'anger'	*meala* 'honey.GEN'	*mealladh* 'enticing'
n^j	n_j	$n_ɣ$	$n^ɣ$
neart 'strength'	*ainm* 'name'	*anam* 'soul'	*nós* 'custom'

Figure 16. Hickey (2011: 406, 407).

It is furthermore known from Irish placenames in the Déise (West Waterford) than /u:/ could be diphthongised to /au/ beyond the cases of vowels before 'tense' sonorants, e.g. *Dún Garbhán* /ˌdaun$_ɣ$ garəˈvɑ:n$_ɣ$/ (= *Dungarvan, WAT*).

Figure 17. Hickey (2011: 362).

(35)		South	West	North
a.	*an-bhreá* 'very fine'	[anˠəvʲrʲɑː]	[anʲvʲrʲɑː]	[ænʲvʲrʲæː]
b.	*seanbhean* 'old woman'	[ʃænˠəvʲænˠ]	[ʃænʲvʲænˠ]	[ʃænʲvʲænˠ]

(69)		Ring	South-Western Irish	
a.	*muintir*	[maɪnʲtʲɪrʲ]	[miːnʲtʲɪrʲ]	'people'
b.	*timpeall*	[haɪmʲpʲəlˠ]	[tʲiːmʲpʲəlˠ]	'around'
c.	*rince*	[raɪnʲkʲə]	[riːnʲkʲə]	'dance'

Sample sentence:

D'ól siad BUIDÉAL *fíona.* 'They drank a bottle of wine.'

Realisations of keyword:

North: [bʷɪdʲɛlˠ] West: [bʷɪdʲeːlˠ] South: [bɪˈd̲eːlˠ]

Sample sentence:

Tá aithne acu ar A HINÍON. 'They know her daughter.'

Realisations of keyword:

North: [ə nʲiːnˠ][152] West: [ə hɪnʲiːnˠ] South: [ə hɪˈnʲiːnˠ]

Figure 18. Hickey (2011: 156, 205, 209, 219) contrast gamma and j sub- and superscript.

(74)	a.	*balla*	[bal$^{\gamma}$ə]	'wall'
	b.	*boladh*	[bal$_{\gamma}$ə]	'smell'
	c.	*buille*	[bɪl^{j}ə]	'blow'
	d.	*buile*	[bɪl$_{j}$ə]	'fury'

In word-final position the low-polarity items do not contrast in polarity with the preceding vowel, that is there are no cases like **m′ʌn′** [m^{j}ʌn$_{j}$] vs. **m′in** [m^{j}ɪn$_{\gamma}$]. A real minimal pair like *mion* [m^{j}ʌn$_{\gamma}$] 'small, fine' versus *mionn* [m^{j}ʌn$^{\gamma}$] 'oath; curse' has the contrast of [n$_{\gamma}$] with [n$^{\gamma}$], but both are preceded by the back vowel [ʌ]. A polarity contrast between consonant +

Figure 19. Hickey (2011: 224).

[n$^{\gamma}$]	The body of the tongue is arched downwards away from the palate; the tip of the tongue is behind the upper teeth (concave tongue configuration).
[n$_{\gamma}$]	There is apico-alveolar contact with slight lowering of the body of the tongue away from the palate.

Figure 20. 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 [n^{j}] (palatal nasal) and [n$_{j}$] (palatalized alveolar nasal).

velarization. Figure 31.3. [d] Figure 31.4. [d$_{\gamma}$] or [d$^{\gamma}$]

- simultaneous release t$_{\gamma}$
- off-glide t$^{\gamma}$

The technical names and articulators of both [d$_{\gamma}$] and [d$^{\gamma}$] are "voiced alveolar velarized stops," and their

Figure 21. Bickford & Floyd (2006: 162). The distinction is simultaneous coarticulation (subscript) vs off-glide (superscript).

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:** *Subscript w y z and ɣ*
2. Requester's name: *Kirk Miller*
3. Requester type (Member body/Liaison/Individual contribution): *individual*
4. Submission date: *2024 October 18*
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: *Superscripts and Subscripts, Latin Extended-G*
2. Number of characters in proposal: *4*
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? *yes*
   If YES explain *Requests for Latin subscripts, with less evidence, were made in L2/21-043 and 21-207R.*
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)? *no*
8. Can any of the proposed characters be considered a presentation form of an existing character or character sequence? *no*
   If YES, is a rationale for its inclusion provided?
   If YES, reference:
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: