

Unicode request for modifier psi and omega

Kirk Miller, kirkmiller, gmail.com

2024 June 24

When reviewing a series of recent proposals that otherwise passed review, the SAH recommended that any Greek-derived phonetic symbols be deferred pending a decision by Unicode as to whether they should be identified as being Latin or Greek script. That decision will impact their properties and naming. The Greek-derived modifier letters are gathered here.

The two modifier letters $\langle\psi\rangle$ and $\langle\omega\rangle$ are requested by the IPA, following evaluation by the IPA Alphabets, Charts and Fonts Committee (Nicolaidis 2024). Modifier $\langle\omega\rangle$ pairs with closed-omega $\langle\omega\rangle$ (U+107A4) in Sinological convention (Figure 1) and so might be similarly encoded as Latin; however, most Greek-derived phonetic letters, such as IPA β^{β} and θ^{θ} and Americanist λ , are encoded as Greek.

There is a parallel to the Sinological pairing of Greek ω and Latin ω (U+0277 LATIN SMALL LETTER CLOSED OMEGA) in the Americanist pairing of Greek λ and Latin λ (U+019B LATIN SMALL LETTER LAMBDA WITH STROKE).

Besides Sinological use, modifier $\langle\omega\rangle$ is attested in Americanist usage as a diacritic for labialized consonants (see figures).

There is no baseline Latin psi for the properties of a Latin modifier $\langle\psi\rangle$ to refer to. It is unlikely that a Greek/Latin distinction would need to be made for either.

Characters

Modifier Greek letters

- ψ 1DFF3 MODIFIER LETTER SMALL GREEK PSI.
- ω 1DFF4 MODIFIER LETTER SMALL GREEK OMEGA. Figure 3 *ff*.

Properties

1DFF3;MODIFIER LETTER SMALL GREEK PSI;Lm;0;L;<super> 03C8;;;N;;;;;
 1DFF4;MODIFIER LETTER SMALL GREEK OMEGA;Lm;0;L;<super> 03C9;;;N;;;;;

The script properties of these characters is Greek, similar to U+1D5D MODIFIER LETTER SMALL BETA (β), U+1DBF MODIFIER LETTER SMALL THETA (θ) and U+1D61 MODIFIER LETTER SMALL CHI (χ).

Annotations

Annotations should identify the historical click letter <Ψ> with its modern equivalent <ɀ>:

MODIFIER LETTER SMALL GREEK PSI

→ 107B9 MODIFIER LETTER RETROFLEX CLICK WITH RETROFLEX HOOK

References

- Ben Braithwaite (2003) Syntactic approaches to possessive constructions in Nuuchahnulth. In Brown & Kalmar (eds.) *Papers for the International Conference on Salishan and Neighbouring Languages* 38: 7–22. University of British Columbia Working Papers in Linguistics.
- Clement Doke (1925) An outline of the phonetics of the language of the ǀhũ: Bushman of the North-West Kalahari. *Bantu Studies* 2: 129–166.
- Katerina Nicolaidis (2024) ‘Unicode support for historical and para-IPA letters.’ Letter submitted to the Unicode Technical Committee, 01 January 2024. [L2/24-049](#).
- Jamin Pelkey (2011) *Dialectology as Dialectic: Interpreting Phula Variation*. De Gruyter Mouton.
- John Stonham (1999) *Aspects of Tsishaath Nootka Phonetics and Phonology*. Lincom Europa.
- (2004) *Linguistic theory and complex words: Nuuchahnulth Word Formation*. Palgrave Macmillan.
- Sun Hongkai [孙宏开] & Jiang Di [江荻] (2004) *Miaoxie Zhongguo yuyan shiyong de guoji yinbiao ji fujia fuhao* [描写中国语言使用的国际音标及附加符号: On IPA and supplementary phonetic symbols used in describing the languages of China]. *Minzu Yuwen* 2004 (1): 47–52.

Chart

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

Latin Extended-G

1DF00

1DFFF

	1DF0	1DF1	1DF2	1DF3	1DF4	1DF5	1DF6	1DF7	1DF8	1DF9	1DFA	1DFB	1DFC	1DFD	1DFE	1DFF
0	ƒŋ	ɣ	ɔ̥	ɟ												
1	ɸ	ɮ	ɔ̥	ɟ												
2	ɔ̥	ɔ̥	ɮ	ɮ												
3	ɣ	ɮ	ɮ	ɮ												ψ
4	ɣ	ɮ	ɮ	ɮ												ω
5	ɮ	ɮ	ɮ	ɮ												ɔ̥
6	ɣ	ɮ	ɮ	ɮ												ɮ
7	ɮ	ɮ	ɮ	ɮ												ɮ
8	ɮ	ɮ	ɮ	ɮ												ɮ
9	ɮ	ɮ	ɮ	ɮ												ɮ
A	ɮ	ɮ	ɮ	ɮ												ɮ
B	ɮ	ɮ	ɮ	ɮ												ɮ
C	ɮ	ɮ	ɮ	ɮ												ɮ
D	ɮ	ɮ	ɮ	ɮ												ɮ
E	ɮ	ɮ	ɮ	ɮ												ɮ
F	ɮ	ɮ	ɮ	ɮ												ɮ

Figures

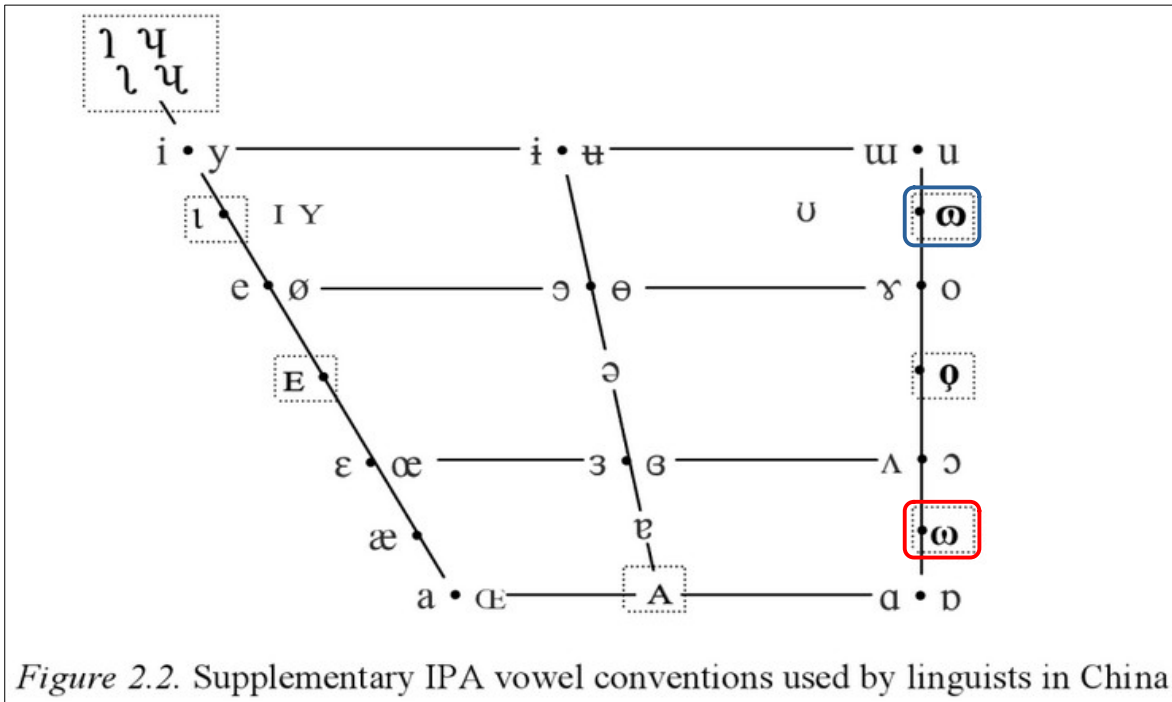


Figure 1. Pelkey (2011: 66). The Sinological para-IPA vowel letters. Omega ω is used alongside U+A7B7 ω LATIN LETTER SMALL CLOSED OMEGA. Modifier $^{\circ}$ is at U+107A4.

表 2: 元音表

舌位前后和唇形 口腔开合 和舌位高低	前		央		后	
	展唇	圆唇	展唇	圆唇	展唇	圆唇
高 (闭)	i	y	i	ɨ	ɯ	u
次高	ɪ (ɿ)	ʏ				ʊ (ω)
中高 (中闭)	e	ø	ə	ɘ	ɤ	o
中	E		ə			Q
中低 (中开)	ɛ	œ	ɜ	ɞ	ʌ	ɔ
次低	æ		ɚ			ω
低 (开)	a	ɶ	A		ɑ	ɒ

Figure 2. Sun & Jiang (2004: 49). The IPA vowel chart, expanded with the Sinological letters <E A Q ω>.

Labialised Stops

Velars and uvulars may be found in labialised varieties, [k^ω] and [q^ω]. This may be an underlying distinction or the result of the presence of an adjacent back rounded vowel, e.g. u + k → uk^ω, a process which will be discussed in greater detail in [7.5.5].

(7)	[k ^ω]	ħa:k ^{ħω} á:λuk ^{ħω}	q ^ħ á:ħ ^á k ^á ač ^ħ iλ
		k ^ħ i:k ^ħ i:sítħin ^á ?ás?i	k ^ħ ik ^ħ iníks ^ħ
	[q ^ω]	q ^{ħω} íyí:ħawu:s:i ^ħ	q ^{ħω} esá:ħ ^á iw ^ó :s:i ^ħ
		q ^ħ á:ʔak ^ħ ?it ^ħ q ^ħ	q ^ħ a: ^ħ

Labialised Fricatives

Only the velar and uvular fricatives are found in labialised forms, as with the stops. Also like the stops, these may be underlying or derived through adjacency with a back rounded vowel.

(10)	[x ^ω]	siχ ^ω it ^ħ	
	[χ ^ω]	níλχ ^ω a:	ʔatχ ^ω ak ^ħ

Figure 5. Stonham (1999: 11–12). Identification of <C^ω> as labialized.

/t^ω/

/t^ω/ is included as a phoneme in Tsishaath because of the phonological processes which certain forms containing [ʔ] undergo, to the exclusion of other forms with the same outward shape. The following morphemes are assigned to the phoneme /t^ω/ rather than to the more common /t/:

čaʔut ^ω	‘rough sea running’
či:t ^ω	‘run away’
-pí:t ^ω	‘...many long objects, songs’
-qu:t ^ω	‘on face’
-ʔut ^ω	‘...place’

/ħ^ω/

For reasons similar to those for /t^ω/ above, /ħ^ω/ is considered a phoneme. Morphemes containing /ħ^ω/ rather than /ħ/ include the following:

čah ^ω	‘cut with adze’
tinħ ^ω	‘calm sea’
?anah ^ω ...ʔis	‘small; size’
?i:ħ ^ω	‘big; great’

Figure 6. Stonham (1999: 27–28). Lateral and pharyngeal fricatives with <ω>.

ISO/IEC JTC 1/SC 2/WG 2
PROPOSAL SUMMARY FORM TO ACCOMPANY SUBMISSIONS
FOR ADDITIONS TO THE REPERTOIRE OF ISO/IEC 10646¹.

Please fill all the sections A, B and C below.

Please read Principles and Procedures Document (P & P) from 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:	<i>Modifier psi and omega</i>
2. Requester's name:	<i>Kirk Miller</i>
3. Requester type (Member body/Liaison/Individual contribution):	<i>individual</i>
4. Submission date:	<i>2024 June 24</i>
5. Requester's reference (if applicable):	
6. Choose one of the following:	
This is a complete proposal:	<input checked="" type="checkbox"/>
(or) More information will be provided later:	<input type="checkbox"/>

B. Technical - General

1. Choose one of the following:		
a. This proposal is for a new script (set of characters):		
Proposed name of script:		
b. The proposal is for addition of character(s) to an existing block:	<input checked="" type="checkbox"/>	
Name of the existing block:	<i>Latin Extended-G</i>	
2. Number of characters in proposal:	<i>2</i>	
3. Proposed category (select one from below - see section 2.2 of P&P document):		
A-Contemporary <input checked="" type="checkbox"/>	B.1-Specialized (small collection) <input type="checkbox"/>	B.2-Specialized (large collection) <input type="checkbox"/>
C-Major extinct <input type="checkbox"/>	D-Attested extinct <input type="checkbox"/>	E-Minor extinct <input type="checkbox"/>
F-Archaic Hieroglyphic or Ideographic <input type="checkbox"/>	G-Obscure or questionable usage symbols <input type="checkbox"/>	
4. Is a repertoire including character names provided?	<input checked="" type="checkbox"/>	
a. If YES, are the names in accordance with the "character naming guidelines" in Annex L of P&P document?	<input checked="" type="checkbox"/>	
b. Are the character shapes attached in a legible form suitable for review?	<input checked="" type="checkbox"/>	
5. Fonts related:		
a. Who will provide the appropriate computerized font to the Project Editor of 10646 for publishing the standard?	<i>Kirk Miller</i>	
b. Identify the party granting a license for use of the font by the editors (include address, e-mail, ftp-site, etc.):	<i>SIL (Gentium Release)</i>	
6. References:		
a. Are references (to other character sets, dictionaries, descriptive texts etc.) provided?	<input checked="" type="checkbox"/>	
b. Are published examples of use (such as samples from newspapers, magazines, or other sources) of proposed characters attached?	<input checked="" type="checkbox"/>	
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)?	<input checked="" type="checkbox"/>	

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? If YES explain	<i>yes</i> <i>deferred pending decision on whether they should be identified as Latin or Greek</i>
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.)? If YES, with whom? If YES, available relevant documents:	 <i>author is a member of the user community</i> -----
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) Reference:	<i>phonetic</i> -----
5. Are the proposed characters in current use by the user community? If YES, where? Reference:	<i>yes</i> <i>see illustrations</i>
6. After giving due considerations to the principles in the P&P document must the proposed characters be entirely in the BMP? If YES, is a rationale provided? If YES, reference:	<i>no</i> ----- -----
7. Should the proposed characters be kept together in a contiguous range (rather than being scattered)?	<i>yes</i>
8. Can any of the proposed characters be considered a presentation form of an existing character or character sequence? If YES, is a rationale for its inclusion provided? If YES, reference:	<i>no</i> ----- -----
9. Can any of the proposed characters be encoded using a composed character sequence of either existing characters or other proposed characters? If YES, is a rationale for its inclusion provided? If YES, reference:	<i>no</i> ----- -----
10. Can any of the proposed character(s) be considered to be similar (in appearance or function) to, or could be confused with, an existing character? If YES, is a rationale for its inclusion provided? If YES, reference:	<i>no</i> ----- -----
11. Does the proposal include use of combining characters and/or use of composite sequences? If YES, is a rationale for such use provided? If YES, reference:	<i>no</i> ----- -----
Is a list of composite sequences and their corresponding glyph images (graphic symbols) provided? If YES, reference:	<i>no</i> -----
12. Does the proposal contain characters with any special properties such as control function or similar semantics? If YES, describe in detail (include attachment if necessary)	<i>no</i> ----- -----
13. Does the proposal contain any Ideographic compatibility characters? If YES, are the equivalent corresponding unified ideographic characters identified? If YES, reference:	<i>no</i> ----- -----