

Unicode request for modifier small capital P

Kirk Miller, kirkmiller, gmail

Denis Moyogo Jacquerye, moyogo, gmail.com

2024 October 18

This request is for modifier small capital P. <P> is a common convention for a voiceless labial trill, IPA [ɸ]. The modifier letter is used for a trilled release, e.g. <p^P> and <b^P> for IPA <p^ɸ, b^ɸ>. These are all small capitals.

Characters

Modifier small capital

^P 1DFD1 MODIFIER LETTER SMALL CAPITAL P.

Properties

1DFD1;MODIFIER LETTER SMALL CAPITAL P;Lm;0;L;<super> 1D18;;;;N;;;;;

References

- Terry Crowley & John Lynch (2006) *The Avava Language of Central Malakula (Vanuatu)*. Pacific Linguistics, The Australian National University.
- Kenneth Olson (2022) The nonexistence of the plain bilabial trill phoneme. *Proceedings of the Linguistic Society of America* 7(1) (<https://doi.org/10.3765/plsa>).

Chart

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

Latin Extended-G

1DF00

1DFFF

	1DF0	1DF1	1DF2	1DF3	1DF4	1DF5	1DF6	1DF7	1DF8	1DF9	1DFA	1DFB	1DFC	1DFD	1DFE	1DFF
0	ḡ	ḥ	ḑ	Ḓ										Ḕ	ḕ	Ḗ
1	ḗ	ḙ	ḑ̇	Ḓ̇										Ḕ̇	ḕ̇	Ḗ̇
2	ḡ̃	ḥ̃	ḑ̃	Ḓ̃										Ḕ̃	ḕ̃	Ḗ̃
3	ḡ̄	ḥ̄	ḑ̄	Ḓ̄										Ḕ̄	ḕ̄	Ḗ̄
4	ḡ̅	ḥ̅	ḑ̅	Ḓ̅										Ḕ̅	ḕ̅	Ḗ̅
5	ḡ̆	ḥ̆	ḑ̆	Ḓ̆										Ḕ̆	ḕ̆	Ḗ̆
6	ḡ̇	ḥ̇	ḑ̇	Ḓ̇										Ḕ̇	ḕ̇	Ḗ̇
7	ḡ̈	ḥ̈	ḑ̈	Ḓ̈										Ḕ̈	ḕ̈	Ḗ̈
8	ḡ̉	ḥ̉	ḑ̉	Ḓ̉										Ḕ̉	ḕ̉	Ḗ̉
9	ḡ̊	ḥ̊	ḑ̊	Ḓ̊										Ḕ̊	ḕ̊	Ḗ̊
A	ḡ̋	ḥ̋	ḑ̋	Ḓ̋										Ḕ̋	ḕ̋	Ḗ̋
B	ḡ̌	ḥ̌	ḑ̌	Ḓ̌										Ḕ̌	ḕ̌	Ḗ̌
C	ḡ̍	ḥ̍	ḑ̍	Ḓ̍										Ḕ̍	ḕ̍	Ḗ̍
D	ḡ̎	ḥ̎	ḑ̎	Ḓ̎										Ḕ̎	ḕ̎	Ḗ̎
E	ḡ̏	ḥ̏	ḑ̏	Ḓ̏										Ḕ̏	ḕ̏	Ḗ̏
F	ḡ̐	ḥ̐	ḑ̐	Ḓ̐										Ḕ̐	ḕ̐	Ḗ̐

Figures

Superscript small-capital <^p> is used for a voiceless bilabial trilled release, analogous to <^b> for voiced and to <^r> for an alveolar trilled release. All three symbols have been used for Avava (Vanuatu). <^p> is equivalent to Americanist superscript *p* with tilde <^{p̃}> in Everett & Kern (1997).

Phonology 31

/aBun/	[a ^m b̃ ^B un]	‘its smell’
/Bel/	[^m b̃ ^B el]	‘(s)he will come’
/Bih/	[^m b̃ ^B ih]	‘how many (irrealis)’
/suBat/	[su ^m b̃ ^B at]	‘coral’

This segment is also found word-finally after a variety of vowels, in which case its bilabial trill release is at least partially devoiced. Thus:

/viaB/	[fia ^m b̃ ^p]	‘Fijian asparagus’
/eneB/	[ene ^m b̃ ^p]	‘deep sea’
/iDoB/	[i ⁿ d ^r o ^m b̃ ^p]	‘it rotted in the ground’

Figure 1. Crowley & Lynch (2006: 31). A voiced bilabial trilled release in [m̃^B] is circled in red, its partially devoiced word-final allophone [m̃^p] in orange. (Also seen in the bottom line is alveolar [n^dr].) On p. 30 the consonant /B/ is described as a “prenasalized voiced bilabial stop with a bilabial trilled release, which typically involves some rounding – or at least thrusting forward – of the lips.” The tilde on <b̃> indicates that secondary labial articulation. The language is Avava of Vanuatu (ISO: [tmb]).

The voiceless stops /p̃/ and /t/ exhibit little notable allophonic variation, being realised invariably as voiceless lightly aspirated labio-velar and alveolar stops respectively. The segment /p/ is generally realised as a voiceless lightly aspirated bilabial stop, though when the following vowel is /u/ and the next syllable begins with /r/, this alternates occasionally with a stop that is released with a voiceless bilabial trill. Thus:

/ipura/	[ipura ~ ip ^p ura]	‘(s)he spat’
---------	-------------------------------	--------------

Figure 2. Crowley & Lynch (2006: 28). Modifier small capital *p* transcribing a voiceless trilled release [p^p] in Avava.

Given the gaps in the inventory, and given the phonetic similarity of affricates and stop-trill sequences, the “Affricate” and “Prestopped trill” rows could be conflated, yielding a more symmetric inventory.

Taking all this into consideration, I would argue that crosslinguistically bilabial stop-trill phonemes pattern as obstruents, and that they are best construed as stops with a bilabial trill release. A transcription that makes use of superscript characters makes this clearer, e.g., [p^ʰ, b^ʙ, ^mb^ʙ].

Figure 3. Olson (2022: 10). Phonemic /pʰ/ and /mbʙ/ in the Vanuatuan language Axamb [ISO: ahb] are realized phonetically as [p^ʰ] and [^mb^ʙ], with trilled releases.

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 small capital P</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 October 18</i>
5. Requester's reference (if applicable):	
6. Choose one of the following:	
This is a complete proposal:	<i>yes</i>
(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):	<i>no</i>	
Proposed name of script:		
b. The proposal is for addition of character(s) to an existing block:	<i>yes</i>	
Name of the existing block:	<i>Latin Extended-G</i>	
2. Number of characters in proposal:	<i>1</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 checked="" 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?	<i>yes</i>	
a. If YES, are the names in accordance with the "character naming guidelines" in Annex L of P&P document?	<i>yes</i>	
b. Are the character shapes attached in a legible form suitable for review?	<i>yes</i>	
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 open font license</i>	
6. References:		
a. Are references (to other character sets, dictionaries, descriptive texts etc.) provided?	<i>yes</i>	
b. Are published examples of use (such as samples from newspapers, magazines, or other sources) of proposed characters attached?	<i>yes</i>	
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)?	<i>no</i>	

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>no</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>proposer is member</i>
3. Information on the user community for the proposed characters (for example: size, demographics, information technology use, or publishing use) is included? Reference:	<i>yes</i>
4. The context of use for the proposed characters (type of use; common or rare) Reference:	<i>linguistics, transcription</i>
5. Are the proposed characters in current use by the user community? If YES, where? Reference:	<i>yes</i> <i>See figures</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>no</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: 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>