Proposal to add two Latin pharyngeal voiced fricative characters

Denis Moyogo Jacquerye <moyogo@gmail.com>
2023-07-18

The Latin characters and casing pair ‘̩’ LATIN CAPITAL LETTER PHARYNGEAL VOICED FRICATIVE and ‘̩’ LATIN SMALL LETTER PHARYNGEAL VOICED FRICATIVE were proposed in L2/21-206, Proposal to encode additional Latin Letters for languages of Argentina, Canada, and the USA, as distinct characters from ‘ʕ’ U+0295 LATIN LETTER PHARYNGEAL VOICED FRICATIVE.

The comments from the Script Ad Hoc were that [the] line arguments in L2/05-194R [for the addition of ‘̩’ U+0242 LATIN CAPITAL LETTER GLOTTAL STOP and a case pairing with ‘ʔ’ U+0294 LATIN LETTER GLOTTAL STOP], may not be applicable. In L2/05-194R the issue was that there are both uni-cameral and bi-cameral writing systems using glottal stop characters, but using U+0294 (gc = Ll) for both would be problematic because the glyph for the uni-cameral usage would more closely resemble the capital letter of bi-cameral usage, not the lowercase letter.

L2/21-206 showed that the practical Pilagá orthography uses a bicameral letter pharyngeal voiced fricative. The unicaleral U+0295 cannot be used because that orthography uses both an uppercase at capital-height and a lowercase at x-height, for example when non-Unicode fonts specific to Pilagá are used, and other writing systems use that character as a unicaleral letter. The writing systems using unicaleral U+0295 are some of those that use the unicaleral U+0294 LATIN LETTER GLOTTAL STOP and are shown in L2/05-194. Additional examples are provided in this document.

Furthermore, some writing systems do use the unicaleral U+0294 LATIN LETTER GLOTTAL STOP as a lowercase of the bicameral capital U+0241 LATIN CAPITAL LETTER GLOTTAL STOP. Some users would like to use a capital character in a similar manner for U+0295 LATIN LETTER PHARYNGEAL VOICED FRICATIVE.

Note: The Unicode 1.0 name of U+0295 was LATIN LETTER REVERSED GLOTTAL STOP. The modifier letter of ‘̩’ U+02E4 MODIFIER LETTER SMALL REVERSED GLOTTAL STOP should not change.
Bicameral uses

The bicameral letter used in Pilagá orthography has the capital at capital-height and the lowercase at x-height. The lowercase is typically not at capital-height nor ascender-height. This orthography has been used by the Pilagá community in the Province of Formosa in Argentina in teaching material published by the Universidad Nacional de Formosa. Additional material using the bicameral letter has been published by the provincial Educación Intercultural Bilingüe or the national Ministry of Education.

Using § U+0295 for both the writing systems using it as a unicameral letter and for the Pilagá writing system using the pharyngeal stop as a bicameral letter is problematic because the glyph of U+0295 resembles the uppercase but is a lowercase character (gc= Ll).

Adding a character as an uppercase of U+0295 would break titlecasing in orthographies using the unicameral U+0295.

| IPA transcription of Pilagá word [taʃadenaʔ] with unicameral U+0295 | taʃadenaʔ |
| Orthographic notation of Pilagá word [taʃadenaʔ] with proposed U+A7CF | taʃadena’ |
| Orthographic all-caps Pilagá word [taʃadenaʔ] with proposed U+A7CE | TAʃADENA’ |

Unicameral uses

The unicameral § U+0295 LATIN LETTER PHARYNGEAL VOICED FRICATIVE is used in orthographies of Nuuchahnulth [nuk] nuučaan̓uł (Nootka in L2/05-194R), Ditidaht [dtd] ditidaqiic̓aq cicqiʔ, Thompson [thp] nteʔkepmxcin but also Okanagan [oka] nsyilxcən, and Columbia-Wenatchi [col] n̓xaʔamxcín or a few others do a lesser apparent degree. It is also used as a unicameral letter in transliterations and phonetic transcriptions.

These writing systems are typically caseless, but have casing in some documents. In those documents the glyph of § U+0295 LATIN LETTER PHARYNGEAL VOICED FRICATIVE has ascender-height, resembling a capital letter, and is the character used for lowercase, all-capitals and titlecase.

In some languages, in titlecased proper nouns, the next letter can be uppercased.

In those systems, the use of § U+0295 mirrors the use of the unicameral ? U+0294 LATIN LETTER GLOTTAL STOP.
<table>
<thead>
<tr>
<th>Orthographic notation of Nuuchahnulth word [ˈeniːtɬ] with unicameral U+0295.</th>
<th>ŋeniiȿ</th>
</tr>
</thead>
<tbody>
<tr>
<td>Atypical all-caps orthographic notation of Nuuchahnulth word [ˈeniːtɬ] with unicameral U+0295.</td>
<td>ŋENIIȿ</td>
</tr>
<tr>
<td>Titlecase transcription of Arabic name [ˈabaːdi] with unicameral U+0295 and following uppercase.</td>
<td>ŋAbādi</td>
</tr>
</tbody>
</table>

### Mixed uses of unicameral and bicameral characters

Some writing systems use or would use both the unicameral ؇ U+0294 and ʕ U+0295 as lowercase and ؇ U+0295 and the proposed U+A7CE as their uppercase, despite them not being in a defined casing pair in Unicode.

For example the Alphabet scientifique des langues du Gabon uses ؇ U+0294 and ؇ U+0295 in a casing pair and text in Shiwa [glottolog: shiw1234] ʃiwé or some dialects of Fang [ISO 639-3: fan] like Fang Meke mə́kə́ may use the letter in bicameral text with that casing pair.

Similarly, some Arabic transcriptions that use U+0294 and U+0295 would use a proposed capital character along with the existing capital U+0241 in capitalization of all-caps or in titlecase.

<table>
<thead>
<tr>
<th>Transcription of Arabic name in titlecase with ؇ U+02BE and ʿ U+02BF in initial position and the following letter uppercased.</th>
<th>ʿAbū ʿUbayd</th>
</tr>
</thead>
<tbody>
<tr>
<td>Transcription of Arabic name in titlecase with ؇ U+0294 and ʕ U+0295 in initial position and the following letter uppercased.</td>
<td>ʿAbū ʿUbayd</td>
</tr>
<tr>
<td>Transcription of Arabic name in titlecase with ؇ U+0241 and proposed U+A7CE in initial position and uppercased.</td>
<td>ʿabū ʿubayd</td>
</tr>
</tbody>
</table>
Proposed additions and changes

UnicodeData.txt
The General category of U+0295 can be changed from Ll to Lo as the General category of U+0294 which was changed from Ll to Lo in Unicode 5.0.

0295;LATIN LETTER PHARYNGEAL VOICED FRICATIVE;Lo;0;L;;;;;N;LATIN LETTER REVERSED GLOTTAL STOP;;;
A7CE;LATIN CAPITAL LETTER PHARYNGEAL VOICED FRICATIVE;Lu;0;L;;;;;N;A7CF;
A7CF;LATIN SMALL LETTER PHARYNGEAL VOICED FRICATIVE;Li;0;L;;;;;N;A7CE;

NameList.txt and Unicode charts
Notes for 0295 can be similar to the notes for 0294.

0295  LATIN LETTER PHARYNGEAL VOICED FRICATIVE
= reversed glottal stop
* this is a caseless letter
* used in IPA, other phonetic notations, and those orthographies which use a caseless glottal stop
* voiced pharyngeal fricative
* ain
  x (latin capital letter pharyngeal voiced fricative - 10XXX1)
  x (latin small letter ezh reversed - 01B9)
  x (modifier letter reversed glottal stop - 02C1)
A7CE  LATIN CAPITAL LETTER PHARYNGEAL VOICED FRICATIVE
A7CF  LATIN SMALL LETTER PHARYNGEAL VOICED FRICATIVE
  x casing use in Pilagá
  x (latin letter pharyngeal voiced fricative - 0295)

Ordering (allkeys.txt)
The ordering of the additional characters can copy the ordering of ʔ 0294 < ɂ 0242 <<< ʔ 0241 < ʰ 02C0:
ʕ 0295 < gne A7CE <<< ɟe A7CF <  ErrorCode
Examples

Figure 1: Vidal, Almeida & Miranda 2014 vol. 2, p. 95 showing LATIN CAPITAL LETTER PHARYNGEAL VOICED FRICATIVE in all-caps text with cap-height and LATIN SMALL LETTER PHARYNGEAL VOICED FRICATIVE in titlecase text with x-height.

Figure 2: Aranda & Centurión 2013, p. 36 showing LATIN CAPITAL LETTER PHARYNGEAL VOICED FRICATIVE in all-caps text with cap-height and LATIN SMALL LETTER PHARYNGEAL VOICED FRICATIVE as a lowercase with x-height in normal text.

Figure 3: Argentina, Ministerio de Educación de la Nación 2015, p. 29 showing LATIN SMALL LETTER PHARYNGEAL VOICED FRICATIVE as a lowercase with x-height.

Pilagá

54  ✮ Grupo familiar pilagá / Ena lawatonawanačak en pitača le’ek
    Ismael Fernández

56  ✮ Nuestrors abuelos / So qadapidi
    Silvia Gricelda Chichole

58  ✮ El hombre mariscador / Ena siača epičačik
    Ismael Fernández
(12) Di’ Paiha wana sa-teto-n ka’lio’ wasač’e malasa hayem
CL Paiha NEG A.1-conocer-ASP ADV conj ADV PRO.1
n’asolek
niño
‘A Paiha no lo conocí en aquel momento porque yo era todavía un niño.’

Además, pueden funcionar como aposiciones. En estos casos (13)–(14), el nombre de persona no está entonacionalmente separado del núcleo y tiene orden fijo dentro de la frase nominal (postnuclear):

(13) so’ Mala’ w/o di’ nibiedtew’o naca hañ’i l-at’e
CL Mala’ EXIST CL B.3-ir-ASP-DIR PT CL.FEM POS.3-madre
Atamana
Atamana
‘A veces Mala’ va (a la casa de) su madre, Atamana.’

Figure 4: Vidal 2015, p. 60, showing LATIN SMALL LETTER PHARYNGEAL VOICED FRICATIVE as a lowercase with x-height.

Da abiamqačeñe nat’ašen
ena’am da w’ó ega adapt’en’á,
nač’e ena’am tašen da am
goitawan.

Figure 5: Quino 2020, showing LATIN CAPITAL LETTER PHARYNGEAL VOICED FRICATIVE in all-caps with cap-height and LATIN SMALL LETTER PHARYNGEAL VOICED FRICATIVE in lowercase with x-height.
Figure 6: Powell 1991, p. 2 showing LATIN LETTER PHARYNGEAL VOICED FRICATIVE and LATIN LETTER GLOTTAL STOP used as unicameral letters with ascender-height in titlecased context in Nuuchahnulth. Nuuchahnulth orthography generally does not use titlecase.

Figure 7: Qʷi[y̓mátkʷ et al. 2014, p. 1 showing LATIN LETTER PHARYNGEAL VOICED FRICATIVE as a unicameral letter with ascender-height in titlecased proper noun in Columbia-Wenatchi. Columbia-Wenatchi orthography generally does not use titlecase except for proper nouns. Note the following letter is uppercased when in initial position in QwÁhn. 
diasect are generally more conservative and homogeneous than sedentary dialects (Watson, 2002: 9). In Jordan, there are many spoken dialects such as ḌAjārmah (Palva, 1976), Bani Ṣaṣar (henceforth BS) (Palva, 1980), Bani Ḥassan Arabic (henceforth BHA) (Irshied, 1984), the Ḥwētāt (Palva, 1986), ḌAbbādi Arabic (Sakarna, 1999), Bani Kenānah (Aldamen, 2007), and Maṣāni Arabic (Rakhia, 2009). Like these dialects, Wadi Ramm Arabic (henceforth WR Arabic), a conservative and homogenous Jordanian dialect, exhibits a large number of linguistic features which distinguish it from other dialects. These differences may be attributed

cult of saints is highly developed in Morocco and undoubtedly was so even before the introduction of Islam, which found itself obliged to tolerate it.⁴ One of the most splendid among all the shrines of Chaouen is that of Mūlāy Ḍil B. r-Rāšīd (= al-Ṣarīf ?Abū l-Ḥasan ?Ali Ben ?Mūsa Ben Rāšīd 1471–1511).⁵ (See the poem 15, “qṣīda d-mūḥāmmad ṣ-ṣrīf l-ṭālamī / Mūḥāmmad ṣ-Ṣrīf l-Ṭālamī’s Poem” the composition of which the man narrator attributed to Mūḥāmmad ṣ-Ṣrīf l-Ṭālamī). He was the founder of the Ben Rāšīd Emirate, and of Chaouen in 1471. Despite his formidable personality and his historical importance, docu-

---

[Figures 8, 9, 10: Descriptions and citations as per the original text]
in word-initial position in titlecased proper nouns in Arabic transliteration. Note the following letter is uppercased for the titlecase in ʻAli or ʻĀlāmi.

Hanitsch, Melanie, 2011. “Kongruenzvariation beim unbelebten Plural im Neu-

Figure 11: Ritt-Benmimoun 2017, p. 285 showing LATIN LETTER PHARYNGEAL VOICED FRICATIVE as a unicameral letter with ascender-height in word-initial position in titlecased proper nouns in Arabic transliteration. Note the following letter is uppercased for the titlecase in ʻAl-ʻArabiyya.

Šahiḥ al-Buxārī = šahīḥ al-Buxārī, accessed through http://sunnah.com, citing the title and number of the kitāb as cited in Wensick (1927) followed by the name of the bāb.
̣abetes Hayyān = al-Baḥr al-Muhīt (ʻAbū Ḥayyān 2010)
̣abetes Ubayd (Faḍā’il al-Qurʾān) = Faḍā’il al-Qurʾān (ʻAbū Ubayd 1995)
Al-Ẓafāš (Maṣāni) = Maṣāni al-Qurʾān (al-Ẓafāš al-‘Awṣat 1995)

Figure 12: van Putten 2022, p. xx showing LATIN LETTER PHARYNGEAL VOICED FRICATIVE and LATIN LETTER GLOTTAL STOP as unicameral letters in titlecased proper nouns in Arabic transliteration. Note the following letter is not uppercased for the titlecase, compare with alternate transliteration shown with the MODIFIER LETTER LEFT HALF RING and MODIFIER LETTER RIGHT HALF RING and the following uppercased letter. Van Putten confirmed in communication the capital would be used if it were available.

Figure 13: Khalafallah 1969, inside cover page showing a form of LATIN CAPITAL LETTER PHARYNGEAL VOICED FRICATIVE used in all-caps as a capital of LATIN LETTER PHARYNGEAL VOICED FRICATIVE in Arabic transliteration. Note that the lowercase form is shown in the following figure.

A DESCRIPTIVE GRAMMAR
OF SĀʿE:DI
EGYPTIAN COLLOQUIAL ARABIC

Figure 13: Khalafallah 1969, inside cover page showing a form of LATIN CAPITAL LETTER PHARYNGEAL VOICED FRICATIVE used in all-caps as a capital of LATIN LETTER PHARYNGEAL VOICED FRICATIVE in Arabic transliteration. Note that the lowercase form is shown in the following figure.
PART I: PHONEMICS

Introduction .......................................................... 11
1. The Phonemes of sâjâdi ........................................... 13
2. Suprasegmentals ..................................................... 15
3. sâjâdi Vocalic System ............................................. 21

Figure 14: Khalafallah 1969, p. 7 showing LATIN LETTER PHARYNGEAL VOICED FRICATIVE in Arabic transliteration. Note that the author’s capital form is shown in the previous figure.

Figure 15: Alihanga 1990, p. 113, showing mixed unicameral LATIN LETTER GLOTTAL STOP and bicameral LATIN CAPITAL LETTER GLOTTAL STOP as a casing pair in the Alphabet scientifique des langues du Gabon.

References


- Alejandra Vidal, “Nombres propios, denominación e identidad entre los pilagá y los wichí (Gran Chaco)” in Bernard Comrie and Lucia Golluscio, *Language Contact and Documentation = Contacto lingüístico y documentación*, De Gruyter, 2015, pp. 51-76. [https://www.academia.edu/48873320/Nombres_propios_denominación_e_identidad_entre_los_pilagá_y_los_wichí_Gran_Chaco](https://www.academia.edu/48873320/Nombres_propios_denominación_e_identidad_entre_los_pilagá_y_los_wichí_Gran_Chaco)
# ISO/IEC JTC 1/SC 2/WG 2
## PROPOSAL SUMMARY FORM TO ACCOMPANY SUBMISSIONS
FOR ADDITIONS TO THE REPERTOIRE OF ISO/IEC 10646

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

Please read Principles and Procedures Document (P & P) from
http://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

See also http://std.dkuug.dk/JTC1/SC2/WG2/docs/roadmaps.html for latest Roadmaps.


### A. Administrative

<table>
<thead>
<tr>
<th>1. Title:</th>
<th>Proposal to add two Latin characters for Pilagá</th>
</tr>
</thead>
<tbody>
<tr>
<td>2. Requester's name:</td>
<td>Denis Moyogo Jacquerye <a href="mailto:moyogo@gmail.com">moyogo@gmail.com</a></td>
</tr>
<tr>
<td>3. Requester type (Member body/Liaison/Individual contribution):</td>
<td>Individual contribution</td>
</tr>
<tr>
<td>4. Submission date:</td>
<td>2023-07-11</td>
</tr>
<tr>
<td>5. Requester's reference (if applicable):</td>
<td></td>
</tr>
<tr>
<td>6. Choose one of the following:</td>
<td>This is a complete proposal: YES (or) More information will be provided later: NO</td>
</tr>
</tbody>
</table>

### B. Technical - General

<table>
<thead>
<tr>
<th>1. Choose one of the following:</th>
</tr>
</thead>
<tbody>
<tr>
<td>a. This proposal is for a new script (set of characters):</td>
</tr>
<tr>
<td>Proposed name of script:</td>
</tr>
<tr>
<td>b. The proposal is for addition of character(s) to an existing block:</td>
</tr>
<tr>
<td>Name of the existing block: Latin Extended-D</td>
</tr>
<tr>
<td>2. Number of characters in proposal:</td>
</tr>
<tr>
<td>3. Proposed category (select one from below - see section 2.2 of P&amp;P document):</td>
</tr>
<tr>
<td>A-Contemporary</td>
</tr>
<tr>
<td>B.1-Specialized (small collection)</td>
</tr>
<tr>
<td>B.2-Specialized (large collection)</td>
</tr>
<tr>
<td>C-Major extinct</td>
</tr>
<tr>
<td>D-Attested extinct</td>
</tr>
<tr>
<td>E-Minor extinct</td>
</tr>
<tr>
<td>F-Archaic Hieroglyphic or Ideographic</td>
</tr>
<tr>
<td>G-Obscure or questionable usage symbols</td>
</tr>
<tr>
<td>4. Is a repertoire including character names provided?</td>
</tr>
<tr>
<td>a. If YES, are the names in accordance with the &quot;character naming guidelines&quot;</td>
</tr>
<tr>
<td>b. Are the character shapes attached in a legible form suitable for review?</td>
</tr>
<tr>
<td>5. Fonts related:</td>
</tr>
<tr>
<td>a. Who will provide the appropriate computerized font to the Project Editor of 10646 for publishing the standard? Denis Moyogo Jacquerye <a href="mailto:moyogo@gmail.com">moyogo@gmail.com</a></td>
</tr>
<tr>
<td>b. Identify the party granting a license for use of the font by the editors (include address, e-mail, ftp-site, etc.):</td>
</tr>
<tr>
<td>6. References:</td>
</tr>
<tr>
<td>a. Are references (to other character sets, dictionaries, descriptive texts etc.) provided?</td>
</tr>
<tr>
<td>b. Are published examples of use (such as samples from newspapers, magazines, or other sources) of proposed characters attached?</td>
</tr>
<tr>
<td>7. Special encoding issue</td>
</tr>
<tr>
<td>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)?</td>
</tr>
</tbody>
</table>

8. 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 http://www.unicode.org for such information on other scripts. Also see UAX#44: http://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.

### C. Technical - Justification
1. Has this proposal for addition of character(s) been submitted before?  
   If YES explain  
   NO

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, available relevant documents:  
   Examples

3. Information on the user community for the proposed characters (for example: size, demographics, information technology use, or publishing use) is included?  
   YES
   Reference: See proposal

4. The context of use for the proposed characters type of use; common or rare)  
   YES
   Reference: See proposal

5. Are the proposed characters in current use by the user community?  
   YES
   If YES, where? Reference: See proposal

6. After giving due considerations to the principles in the P&P document must the proposed characters be entirely in the BMP?  
   YES
   If YES, is a rationale provided?  
   Yes

7. Should the proposed characters be kept together in a contiguous range (rather than being scattered)?  
   YES

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: See proposal

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?  
    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: