Unicode request for two BMP Latin characters | L2/22-113R

Kirk Miller, kirkmiller@gmail.com

2022 July 09

This request is for two Latin letters used in basic literacy material. We request that they be added to the BMP.

The Luiseño community of California (Uto-Aztecan, ISO [ʟᴜɪ]) uses the casing pair ⟨ẞ ẞ⟩ in their literacy and language-revival program. Myra Masiel-Zamora, Curator of Pechanga Cultural Resources in Temecula, has asked that we adopt ⟨ẞ ẞ⟩ into Unicode to aid those efforts. These two letters are also used in descriptive Cupeño material (ISO [ᴄᴜᴘ]).

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

Characters

ẞ   A7CC LATIN CAPITAL LETTER S WITH DIAGONAL STROKE. Figures 1–2.

ẞ   A7CD LATIN SMALL LETTER S WITH DIAGONAL STROKE. Figures 1–3.

Properties

```
A7CC;LATIN CAPITAL LETTER S WITH DIAGONAL STROKE;
    Lu;0;L;;;;;N;;;;A7CD;
A7CD;LATIN SMALL LETTER S WITH DIAGONAL STROKE;
    Ll;0;L;;;;;N;;;A7CC;;A7CC
```

Chart

As they are needed for practical orthographies, the proposed characters should be added to the BMP. Proposed characters are shown on a red background, existing characters on grey. The full Extended-D block is reproduced at the end of this document.

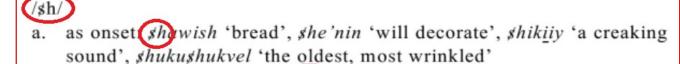|	...0	...1	...2	...3	...4	...5	...6	...7	...8	...9	...A	...B	...C	...D	...E	...F
Latin Extended-D																
U+A7Cx	Ꟁ	ꟁ	Ꟃ	ꟃ	Ᶎ	Ʂ	Ꟈ	Ꟈ	ꟈ	Ꟊ	ꟊ	ʏ	ẞ	ẞ		
U+A7Dx	▢	▢		ꟼ		ꟽ	▢	▢	ꟸ	ꟹ						
U+A7Ex																
U+A7Fx			ꟲ	ꟳ	ꟴ	ꟷ	ꟸ	ꟹ	ʜ	œ	ꟺ	Ꟶ	ꟼ	Ɱ	Ꞁ	ꟿ

References

Raymond Basquez Sr, Neal Ibanez & Myra Masiel-Zamora (2018) *'Atáaxum Alphabet.* Great Oak Press, Pechanga Band of Luiseño Mission Indians.

Jane Hill (2005) *A Grammar of Cupeño.* University of California Publications.

Figures

S, s with slash (Ꞩ, ꞩ)

Used for Luiseño and Cupeño; in the latter it is restricted to the digraph ⟨ꞩh ꞩh⟩ for the retroflex fricative. Luiseño uses ⟨Ꞩ ꞩ⟩ in literacy material and has a current need. These letters are distinct from the existing S with an oblique stroke, U+A7A8 ⟨S⟩ and U+A7A9 ⟨s⟩.

Figure 1. Hill (2005: 14), illustrating the Cupeño digraph ⟨ꞩh ꞩh⟩

Figure 2. Basquez (2018). Contrast of the three Luiseño sibilants ⟨S s⟩, ⟨Ꞩ ꞩ⟩ and ⟨Sh sh⟩.

Figure 3. Basquez (2018). The proposed character ⟨ś⟩ illustrated on the front cover of a Luiseño primer.

Code block chart

Here is the full Latin Extended-D block. The proposed letters are in pink.

A720   **Latin Extended-D block**   **A7FF**

	A72	A73	A74	A75	A76	A77	A78	A79	A7A	A7B	A7C	A7D	A7E	A7F
0	꜠	ꜰ	Ꝁ	Ꝕ	Ꝿ	ꝰ	Ʇ	Ꞑ	Ꞡ	Ʞ	Ꟁ	⯀		ꟲ
1	꜡	ꜱ	ꝁ	ꝕ	ꝡ	ꝱ	ꞁ	ꞑ	ꞡ	Ꞣ	ꟁ	⯀		ꟳ
2	Ꜣ	Ꜳ	Ꝃ	Ꝗ	ꝲ	ꝺ	Ꞃ	Ꞓ	Ꞣ	Ꞵ	Ꟃ			ꟴ
3	ꜣ	ꜳ	ꝃ	ꝗ	ꝳ	Ꝼ	ꞃ	ꞓ	ꞣ	Ꞧ	ꟃ	Ᵽ		Ꞙ
4	Ꜥ	Ꜵ	Ꝅ	Ꝙ	Ꝼ	ꝼ	Ꞅ	Ꞔ	Ꞥ	ꞵ	Ꞓ			Ꞟ
5	ꜥ	ꜵ	ꝅ	ꝙ	Ᵹ	Ᵹ	ꞅ	ꞕ	ꞥ	ꞵ	ꞩ	ꞵ		ꟻ
6	Ꜧ	Ꜷ	Ꝇ	Ꝛ	Ꝿ	Ꝿ	Ꞇ	Ꞗ	Ꞧ	Ꞷ	Ꞵ	⯀		ꟼ
7	ꜧ	ꜷ	ꝇ	ꝛ	ꝿ	ꝿ	ꞇ	ꞗ	ꞧ	ꞷ	Đ	⯀		ꟽ
8	Ꜩ	Ꜹ	Ꝉ	Ꝡ	Ɣ	ꝸ	˄	Ꞙ	Ꞩ	Ʋ	ꞩ	ꞵ		ꟾ
9	ꜩ	ꜹ	ꝉ	ꝡ	ɣ	Ꝺ	:	ꞙ	ꞩ	ꞹ	ꞩ	ꞵ		ꟿ
A	Ꜫ	Ꜻ	Ꝋ	ꝛ	ꜫ	ꝸ	=	Ꝺ	Ɥ	Ꞛ	ꞩ			Ɜ
B	ꜫ	ꜻ	ꝋ	ꝛ	ꜫ	ꝼ	'	ꝺ	Ꞛ	ꞛ	Ɣ			Ⅎ
C	Ꜭ	Ꜽ	Ꝍ	ꝛ	Ꜭ	ꝼ	'	Ꝺ	ꞡ	Ꞌ	𝒮			Ꟑ
D	ꜭ	ꜽ	ꝍ	ꝛ	ꜭ	ꝼ	Ꞑ	ꝺ	Ꞌ	ꞌ	𝒮			Ɯ
E	Ꜯ	Ꜿ	Ꝏ	Ꝝ	ꝯ	ꝼ	ꞓ	ᴚ	Ɪ	ʼꞟ				Ɪ
F	ꜯ	ꜿ	ꝏ	ꝝ	ꝯ	ꝼ	·	ᴚ	Ꞟ	ꞹ				ꟿ

4

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:** *Modifier capital letters (b)*

2. Requester's name: *Kirk Miller*

3. Requester type (Member body/Liaison/Individual contribution): *individual*

4. Submission date: *2022 July 09*

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):

   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-D*

2. Number of characters in proposal: *2*

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 http://www.unicode.org for such information on other scripts. Also see Unicode Character Database (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? _____ *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? _____ *Myra Masiel-Zamora (Curator of Pechanga Cultural Resources)* _____

        If YES, available relevant documents: _____ *p.c. emails available* _____

3. Information on the user community for the proposed characters (for example:

    size, demographics, information technology use, or publishing use) is included? _____ *2,500* _____

    Reference: _____ *2,500 Luiseño: Moseley (2007) Encyclopedia of the world's endangered languages* _____

4. The context of use for the proposed characters (type of use; common or rare) *literacy material*

    Reference: _____

5. Are the proposed characters in current use by the user community? _____ *yes* _____

        If YES, where? Reference: _____ *See references* _____

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* _____

        If YES, reference: _____ *use for basic literacy material* _____

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: _____

9. Can any of the proposed characters be encoded using a composed character sequence of either

    existing characters or other proposed characters? _____ *yes (with U+0338)* _____

        If YES, is a rationale for its inclusion provided? _____ *yes* _____

        If YES, reference: _____ *TUS 14.0, p. 330, "Overlaid diacritics"* _____

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? _____ *yes (U+A7A8, A7A9)* _____

        If YES, is a rationale for its inclusion provided? _____ *yes* _____

        If YES, reference: _____ *SAH review, 2022 May 22* _____

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: _____