Proposal to encode two Latin characters for Mazahua

Denis Moyogo Jacquerye <moyogo@gmail.com>
2016-01-22
(Revision of L2/14-201)

Introduction

The Mazahua language (*jñatrjo* or *jñatjo* in Mazahua, ISO 639-3: maz), a Mesoamerican language recognized by a statutory law in Mexico, is written with the Latin alphabet with additional letters. One of its several Latin orthographies, the practical orthography of the SEP (Secretaría de Educación Pública de México) is used in some bilingual primary schools. The SEP orthography uses the diagonal stroke on nasalized vowels a, e, o and u.

The a with stroke, e with stroke, o with stroke have already encoded characers: Ⱥ ⱥ Ɇ ɇ Ø ø.

- `U+023A LATIN CAPITAL LETTER A WITH STROKE`
- `U+2C65 LATIN SMALL LETTER A WITH STROKE`
- `U+0246 LATIN CAPITAL LETTER E WITH STROKE`
- `U+0247 LATIN SMALL LETTER E WITH STROKE`
- `U+00D8 LATIN CAPITAL LETTER O WITH STROKE`
- `U+00F8 LATIN SMALL LETTER O WITH STROKE`

Some Mazahua documents currently use Ʉ (`U+0055 LATIN CAPITAL LETTER U`, `U+0338 COMBINING LONG SOLIDUS OVERLAY`) and ʉ (`U+0075 LATIN SMALL LETTER U`, `U+0338 COMBINING LONG SOLIDUS OVERLAY`), or `U+0338` combined with other base letters. However there is precedent for the encoding of letter characters with overlay stroke as single characters, for example Ⱥ ⱥ Ɇ ɇ were encoded in 2004.

Proposed characters

| A7B8 | Ʉ | LATIN CAPITAL LETTER U WITH STROKE | • used in Mexico |
| A7B9 | ʉ | LATIN SMALL LETTER U WITH STROKE | |

Examples

UDHR – Mazahua Central

JÑA 'A YO PJOSṶ TEXE YE TE̱'E̱ NUYO KA̱NRA̱ A XESE NE XOÑIJOMU

PJṲRṲ

Panrṵji k'ṵ mara mimiji na joo texe yo nte̱'e̱ in dya ra chṵnji, numa so'o k'ṵ ra mamṵji k'ṵ ri chjekjoji texezgoji.

Panrṵji k'ṵ ma dya ri pjechiji texto yo jña'a nuyo nrra pjosṵ ye nte̱'e̱, nudyama so'o k'ṵ ra potr'ṵji yo in Kjuarma; ngek'ua nesta nura mbaraji k'ṵ ye nte̱'e̱, ma sa̱ja̱ kja ne xoñi jomṵ, na kjuana k'ṵ rambesi nu zi'i, ra so'o mamṵ nu in jña'a ñeje ra eme nu K'inchi angeze.

Figure 1: Example from UDHR - Mazahua Central.

¿Ja ga kjazgojme yo ri testigojme e Jehová?

Ri ka̱xtrjojme kja texe nu Xoñijomu, zo nrre ri n'añozgojme ñe pje nrre idioma ri ñagojme, nrrexezgojme ri nee ra kjagojme k'o mama e Jehová. K'o mama kja ne Biblia nge o Jñaa Mizhokjimi. Angeze go dya̱tr'a̱ texe k'o b'ub'u ajens'ẹ ñe kja ne Xoñijomu. Ri kjagojme k'o go tsjaa e Jesucristo, ngek'ua xitsigojme ri cristianogojme. Ri jichijme yo ntee pje mama ne Biblia ñe ja ga kjaa o Reino e Mizhokjimi. Yo ntee pa̱ra̱ji ke ri testigojme e Jehová nange ri pjøs'ujme yo ntee ra mba̱ra̱ji kjo ngeje e Jehová ñe pje ra tsjaa o Reino.

Ri ñuu pje b'ub'uba kja ne sitio. Søø ri xøruba ne Biblia. Søø ri pjechi pje ri kjagojme ñe kjo ngezgojme.

Figure 2: Example from JW.org.

Ma kja na punkju e jyarṵ ngue ma Abril ñe Mayo.Nudya b'ub`ṵ na punkju e tee ko echi ye saa.Nuyoe saa ko tee ngueje e: oyamel, cedro, tizhi,menza, madroño, xisaa. B`ṵb`ṵ nrre e saa ko tee kja j`ñiji ngueje e pedyṵṵ, tepozan, nrrenze.Yo nrra̱jna̱ ko tee ngueje eucalipto, sabino, fresno, casuarinna,trueno neje axle. Ye xisa ko tee ngueje e baxṵ,jñuxṵ, ñe in pjiño ye rrekjua. ko tumṵ, yo tee ngueje e ixi, mbokjijñi, ts`ana, pjiño, ko saaji, ko ne chuu nzinzana, navo, in tsijñi ye b`aga. Neje tee yo pjiño ko na jo e rrekjua ko ne chuu arnica, pjiño e rrekjua, xumujṵ joopjiño, altamisa, piru, tsiguarṵ, t`oopezhto, gordolobo, mirtho, neje manzanilla.

yo nrra̱jna̱ ko tee ne chuu rroxaxi , alcatraz, clavel, in nrrajna e tsidyo.Yo dyoxṵ ko tee ngueje miño, tr`amixi kjimi, mbexchkua, nrranturṵ, Ṵjmṵ, tizhi, armadillo, tlacuache, mixi treje, suu ko ne chuu, tortola, sṵṵ nrrixu. Bṵbṵ e pjantr`eje neje na punkju e tr`adyoxṵ ko ne chuu e pjadṵ, rrekua, mulas, kuchi, ngoñi, b`aga , nrrenchurṵ, neje napunkju e s`uu ko ne chuu gorrio, ts`i nrrunturṵ, cholona ne s`uu kjo tonjo me na soo.

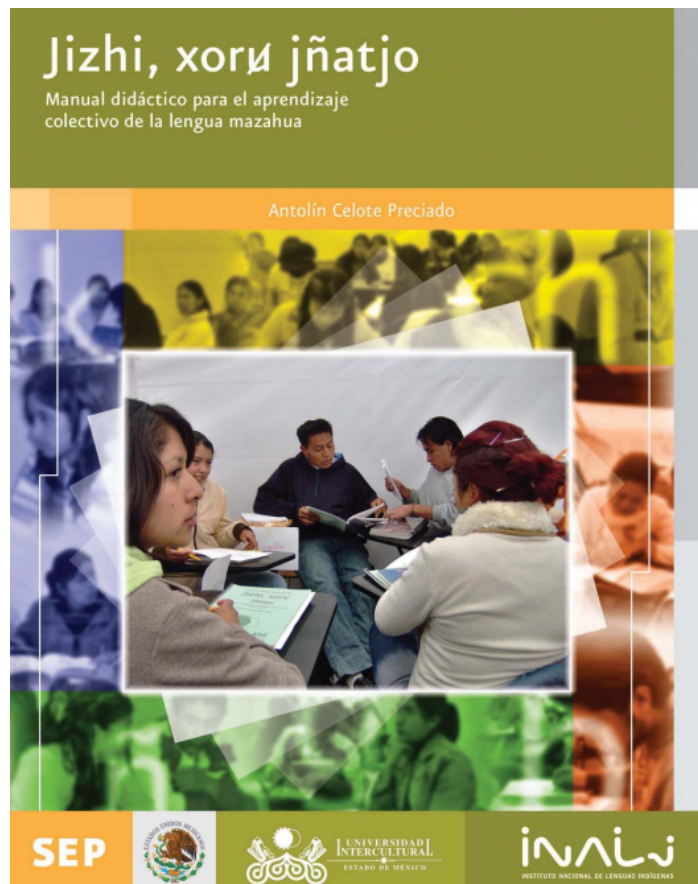Figure 3: Example from cdi.gob.mx.

Figure 4: Example from Celote Preciado 2007 (cover page).

Figure 5: Example from Twitter.

References

Antolín Celote Preciado, *Jizhi, xorɇ jñatjo*, Secretaría de Educación Pública de México (SEP), Instituto nacional de lenguas indígenas, 2007.

Estado de México, *Mazahua de San Felipe del Progreso*, http://www.cdi.gob.mx/mazahuas_edomex/index_m.html

Office of the High Commissioner for Human Rights, *Universal Declaration of Human Rights - Mazahua Central*,
   http://www.unicode.org/udhr/d/udhr_maz.pdf

Oficina de la Presidencia de la República, *Presidencia México @PresidenciaMX*, tweet in Mazahua on Twitter, 2013-10-07,
   https://twitter.com/PresidenciaMX/status/387324007701561344

Watch Tower Bible and Tract Society of Pennsylvania, *Yo testigo e Jehová*, http://www.jw.org/jw-mzh/

Wikipedia, *Idioma mazahua*, https://es.wikipedia.org/wiki/Idioma_mazahua

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

A. Administrative

1. **Title:** *Proposal to encode two Latin characters for Mazahua*

2. Requester's name: *Denis Moyogo Jacquerye*

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

4. Submission date: *2015-01-22*

5. Requester's reference (if applicable):

6. Choose one of the following:

   This is a complete proposal: *Complete*

   (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: *A720-A7FF*

   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 *A*   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?

   Denis Moyogo Jacquerye

---

b. Identify the party granting a license for use of the font by the editors (include address, e-mail, ftp-site, etc.):

Denis Moyogo Jacquerye moyogo@gmail.com

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.

A7B8;LATIN CAPITAL LETTER U WITH STROKE;Lu;0;L;;;;;N;;;;A7B9;
A7B9;LATIN SMALL LETTER U WITH STROKE;Ll;0;L;;;;;N;;;A7B8;;A7B8

C. Technical - Justification

1. Has this proposal for addition of character(s) been submitted before?   *Yes*

   If YES explain   *This is a revised version that takes into account feedback on previous version (L2/14-201)*

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?   *Mary Hopkins (SIL)*

   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?   *Yes*

   Reference: *Educational material, government documents, on Internet*

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

   Reference: *Common within the community*

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

   If YES, where? Reference: *Mexico*

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?   *No*

   If YES, reference:

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*

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

   If YES, reference: *Unique characters for Latin letter + overlay diacritic*

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: