

To: UTC and ISO/IEC JTC1/SC2 WG2
Title: Proposal to encode LISU LETTER YHA in the UCS
From: Lorna Priest Evans (SIL International), Neil Rees (British and Foreign Bible Society)
Date: 16 January 2019

Introduction

We wish to propose the addition of LISU LETTER YHA in the “Lisu (Lisu)” script. The current Lisu block (U+A4D0..U+A4FF) does not have any more room for another character. Thus, a new column will need to be added to the Lisu block. Since the script is not casing, it should be acceptable to encode the new block (Lisu Supplement) in the SMP at the recommended position of U+11FB0. This will require one column for the Lisu script at U+11FB0..U+11FBF.

The character follows the properties of similar characters as described in “Character Name and annotations”.

The Naxi [nxq] language of China is spoken by 300,000 [Ethnologue]. The Naxi language is primarily written using the Latin script. However, it was written (perhaps experimentally) in the 1920s and 1930s using the Fraser i.e. Lisu script [Ethnologue]. Recent efforts to digitize the historical documents have made it necessary to encode this character. In reviewing the original Unicode proposals for Lisu, this character does not appear to have ever been proposed.

In the “Figures” section we have examples from three different documents showing the use of this character. There are 360 instances of LISU LETTER YHA in the Gospel of Mark and 1,062 instances of the “LISU LETTER YA” in the Gospel of Mark. This shows that it is not quite as commonly used as YA but it is also not uncommon. The shapes for the two characters:

YA: Y

YHA: ʌ

This character looks like U+2144 TURNED SANS-SERIF CAPITAL Y (with Sm property). However, for the same reasons that the Lisu script was encoded, this character should be encoded in a Lisu script block with similar properties to other Lisu characters.

Thus, we wish to request the encoding of LISU LETTER YHA with properties as described below.

Character Name and annotations

11FB0	LISU LETTER YHA • Naxi language
-------	------------------------------------

Unicode Character Properties

Properties for the new character:

11FB0; LISU LETTER YHA; Lo; 0; L; ; ; ; ; N; ; ; ;

This character has no decomposition.

This character can sort immediately after:

A4EC ; [.4658.0020.0002] # LISU LETTER YA

Figures

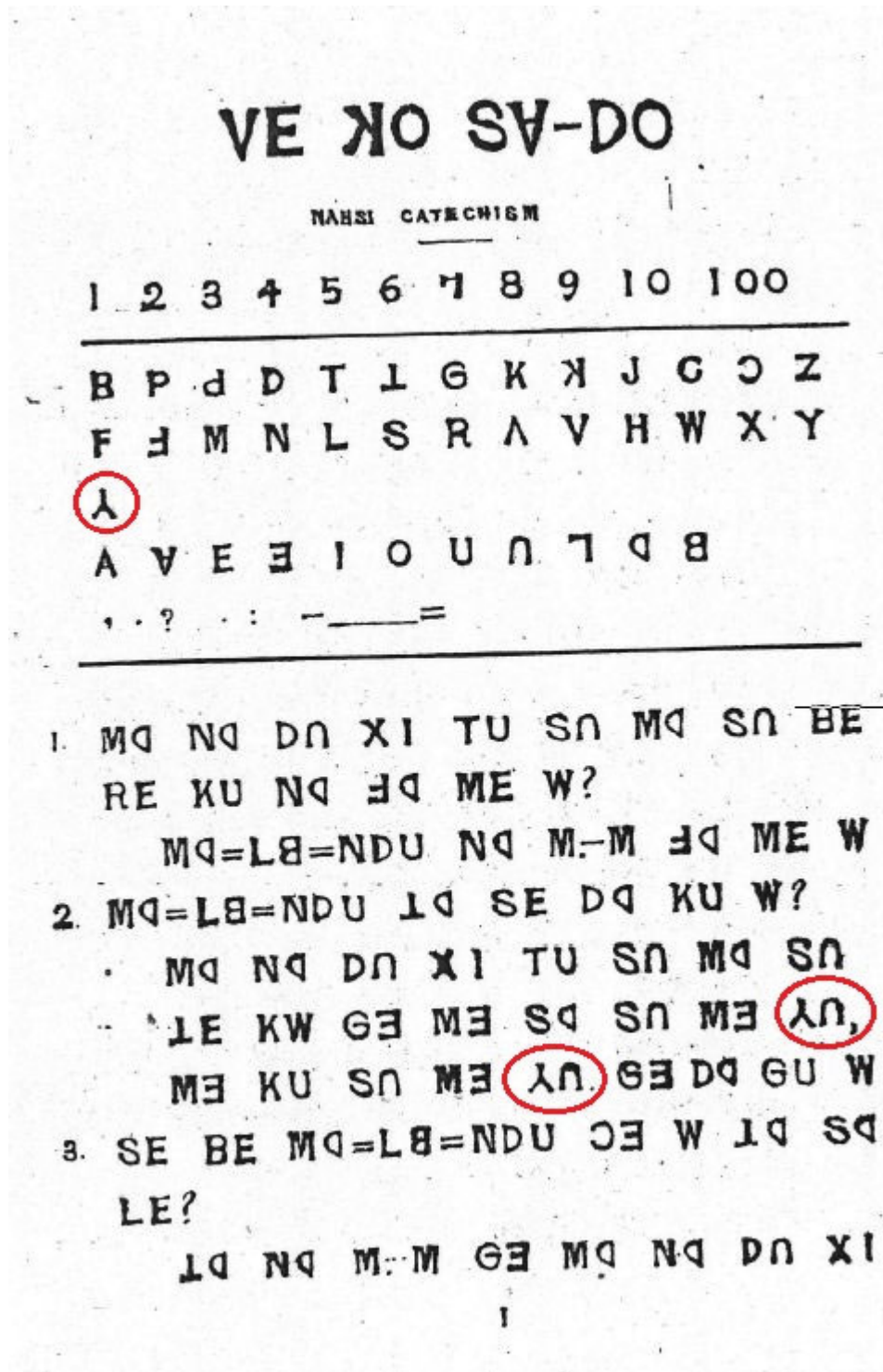


Figure 1. LISU LETTER YHA (Naxi Catechism, p. 1)

VE KU-DD NZL IE-D

NAHSI HYMNBOOK

- 1 YE=SU TO, LO: 82. PH
- 1 XI GE NU-ME HU-CI XU
MD KU BD BV, JU DU RT
CI-CI KV-KV OD XT BE
K ME DD SU, DD ME LO
- 2 NZV NDU KV ND LE-PU-TU
ST NDU KV ND XU, LE NJI
XI NI IE BE PE EYN XU
VE GE HU-CI DD ME LO
- 3 A-NGD O-LU VE GD-YI
DU LO XI CIT ME VV SU
VE=NJT XI CIT KU-KU YE
KW ME LE KU, K ME DD
- 4 NDT UN ME SE, KV-VWE LU
VE=NJT DD SE, NDT LE SD
YE=SU TO IO: JU XT KU
K ME LE DD, NU-ME BI:

Figure 2. LISU LETTER YHA (Naxi Hymnbook, p. 1)

M-KO LE-D

I

YO=V XI CT VE KO XE.

- 1 VE GE RO YE=SU CI-TU K. OK
GE KU. YU.
- 2 VE-XI YI=S=Y GE LE-D LO
LE GIE-TU GE ID NI-NI-W XE
ME LU LE VE ND VE GE CT-
XI CT YE ND GE KV YU, RD-
GU VO-VO CT.
- 3 DU-BV LO ND XI GE KO DD
DD. GU NYU LT-LT. BE XE ME
VE RO GE RD DD LE VO.-VO
RD GU TU-TU BE LE M-M XE Y
- 4 CD RU NGU CU BE YO=V XE
GE DD GU FD YE DU-BV LO ND
XI CT ND KW ME KU-KV GE
CT-CT KU-NDU LE ME YE, FU
LE RU DD CT.

I

Figure 3. LISU LETTER YHA (Naxi Gospel of Mark, p. 1).

References

Scharten Catechism Nahsi 1933

Scharten Hymnbook Nahsi 1933

BFBS. 1932. NA-HSI Mark. British & Foreign Bible Society. Shanghai.

Simons, Gary F. and Charles D. Fennig (eds.). 2018. *Ethnologue: Languages of the World*, Twenty-first edition. Dallas, Texas: SIL International. (<http://www.ethnologue.com> accessed 10-Dec-2018).

The Unicode Consortium. *The Unicode Standard*, Version 11.0.0, (Mountain View, CA: The Unicode Consortium, 2018. ISBN 978-1-936213-19-1) (<http://www.unicode.org/versions/Unicode11.0.0/> accessed 28-Nov-18).

**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 <http://www.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://www.dkuug.dk/JTC1/SC2/WG2/docs/summaryform.html>.

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

A. Administrative

1. Title:	Proposal to encode LISU LETTER YHA in the UCS		
2. Requester's name:	Lorna Priest Evans and Neil Rees		
3. Requester type (Member body/Liaison/Individual contribution):	Individual contribution		
4. Submission date:	10-Dec-2018 (revised 16-Jan-2019)		
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:			No

B. Technical – General

1. Choose one of the following:			
a. This proposal is for a new script (set of characters):			No
Proposed name of script:			
b. The proposal is for addition of character(s) to an existing block:			Yes
Name of the existing block:	Lisu (a new block for Lisu script is needed)		
2. Number of characters in proposal:			1
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"/>
C-Major extinct	<input type="checkbox"/>	D-Attested extinct	<input type="checkbox"/>
F-Archaic Hieroglyphic or Ideographic	<input type="checkbox"/>	E-Minor extinct	<input type="checkbox"/>
G-Obscure or questionable usage symbols	<input type="checkbox"/>		
4. Is a repertoire including character names provided?			
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?	Michael Everson		
b. Identify the party granting a license for use of the font by the editors (include address, e-mail, ftp-site, etc.):			
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)?			

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.

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>Proposal is based on attested usage.</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> <i>See proposal</i>
4. The context of use for the proposed characters (type of use; common or rare) Reference:	<i>Yes</i> <i>See proposal</i>
5. Are the proposed characters in current use by the user community? If YES, where? Reference:	<i>Yes</i> <i>See proposal</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> <i>Yes</i> <i>See proposal</i>
7. Should the proposed characters be kept together in a contiguous range (rather than being scattered)?	<i>n/a</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>Yes</i> <i>Yes</i> <i>See proposal</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> <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>